

2018 年度 修士論文

コンテキスト検索エンジンへの  
論理演算機能導入に関する研究  
**Study on Introduction of Logical Operation  
Context Search Engine**

指導教員 高間 康史 教授  
首都大学東京大学院  
システムデザイン研究科  
システムデザイン専攻  
情報通信システム学域

17890518 佐藤宏貴

2019 年 3 月 12 日

## 目次

|       |                                 |    |
|-------|---------------------------------|----|
| 1     | はじめに                            | 1  |
| 2     | 関連研究                            | 3  |
| 2.1   | 高度検索エンジンに関する研究                  | 3  |
| 2.1.1 | 既存 Web 検索エンジンの問題点               | 3  |
| 2.1.2 | 高度な検索エンジン                       | 4  |
| 2.1.3 | コンテキスト検索エンジン                    | 5  |
| 2.2   | Web 検索エンジンでの情報検索支援              | 11 |
| 2.3   | 検索エンジンでの情報探索行動                  | 13 |
| 3     | コンテキスト検索エンジンでの効率的な絞り込み支援        | 16 |
| 3.1   | コンテキスト検索エンジンへの論理演算機能の導入         | 17 |
| 3.2   | コンテキスト検索エンジンのガイドラインの作成          | 23 |
| 3.3   | コンテキスト検索エンジンでの効率的な検索行動支援        | 24 |
| 4     | 評価実験                            | 26 |
| 4.1   | 論理演算機能を導入したコンテキスト検索エンジンに関する比較実験 | 26 |
| 4.1.1 | 実験概要                            | 26 |
| 4.1.2 | 検索タスクの説明                        | 27 |
| 4.1.3 | 実験の結果・考察                        | 31 |
| 4.2   | 予備実験                            | 46 |
| 4.2.1 | 実験概要                            | 46 |
| 4.2.2 | 実験の結果・考察                        | 46 |
| 4.3   | 効率的な検索行動に対する評価実験                | 52 |
| 4.3.1 | 実験概要                            | 52 |
| 4.3.2 | 実験の結果・考察                        | 52 |
| 5     | おわりに                            | 61 |
|       | 謝辞                              | 62 |
|       | 参考文献                            | 63 |
|       | 発表文献                            | 66 |
|       | 付録 A コンテキスト検索エンジンのガイドライン        | 67 |

|                     |    |
|---------------------|----|
| 付録 B 効率的な検索行動に対する知識 | 80 |
| 付録 C 実験に用いた記事       | 83 |

# 1 はじめに

本論文では，コンテキスト検索エンジンにおける効率的な検索結果の絞り込み支援を目的とした，論理演算機能の導入を提案する．

多くの人が日常的に利用するようになったインターネット上では膨大な量の情報が蓄積され続けている．現在 Web 上の情報へのアクセス手段として，Web 検索エンジンが一般に利用されている．既存の Web 検索エンジンは，ユーザが入力したキーワードと一致した Web ページを検索結果として返すといった直感的でわかりやすい検索機能により幅広いドメインで利用されている．しかし，Web 上の情報が増えていくに従い，ユーザが必要とする情報を大量の情報の中から発見することは難しくなっており，検索エンジンが提供する基本検索機能とユーザの情報要求との乖離が大きくなっていることが指摘されている [14]．また，ユーザの情報要求も多様化してきており，ユーザの情報要求をキーワードと一致した Web ページを返すという方法で満たすのには限界がある．

このような問題点に対し，自然言語によるクエリ入力を受け付ける方法 [6][7] [8] や検索対象を特定のドメインに限定することで，高度な検索機能を提供する専門検索エンジンが提案されている [9][10][11][12]．

ドメインを限定せずに，既存検索エンジンよりも高度な検索機能を提供することで，ユーザの多様な情報要求に答える事を目的とした方法の一つとして，タスクを「動向に関する問い」に限定するコンテキスト検索エンジンが提案されている [14]．コンテキスト検索エンジンでは，時系列データを対象とした検索エンジンであり，時間的観点から関係のあるアイテムを発見することが可能となっている．しかし，独自に定義された構文に従ったクエリの入力が必要であり複雑なクエリ生成に要するユーザの負担増加と，データ量の増大による検索結果の効率的な絞り込みが課題となっている．そこで本論文では，より効率的な検索を実現するために，論理演算機能の導入を目的とする．

既存の Web 検索エンジンでは，キーワードによる検索が行われており，AND，OR，NOT 等の論理演算子や，複数のキーワードの位置関係を指定する近接演算子やフレーズ検索によって検索結果を絞り込むことができる．しかし，既存の Web 検索エンジンと同様の論理演算機能をコンテキスト検索エンジンに導入した場合，クエリの記述がさらに複雑化する問題点がある．そこで提案手法では，クエリの記述に論理演算子を利用するのではなく，検索結果に対し論理演算を適用する．また，一般ユーザの多くは既存の Web 検索エンジンにおいて論理演算子を用いた複雑な検索式を用いず，一つのキーワードのみを用いた単純な検索を行うことが多いことが知られており [23]，この問題点は，コンテキスト検索エンジンでも当てはまることが想定される．そこで本論文では，コンテキスト検索エンジンの利用に関するガイドラインを作成することで，ユーザに対して利用方法についての適切な知識を提供するとともに，効率的な検索行動に対する知識も提示することでコンテキスト検索エンジンの効果的な活用を支援する．

論理演算機能を導入したコンテキスト検索エンジンと従来のコンテキスト検索エンジンで比較実験を行い，論理演算機能の有効性について考察する．また，予備実験から収集した検索行動ログを分析した結果から，コンテキスト検索エンジンにおける効率的な検索行動に対する知識を作成し

た．ユーザ実験を行い，作成したガイドラインや効率的検索行動に関する知識の有効性について考察する．

## 2 関連研究

### 2.1 高度検索エンジンに関する研究

#### 2.1.1 既存 Web 検索エンジンの問題点

現在、通信技術の発達により、幅広く多様な人が利用しているインターネット上では毎日大量の Web ページが作成又は更新、削除されており、膨大な情報が蓄積されている。この膨大な情報源の活用する方法として、Google<sup>\*1</sup>や Yahoo!<sup>\*2</sup>などの Web 検索エンジンを用いて必要な情報を見つける手段が一般的なものになっている。その理由として、既存 Web 検索エンジンはユーザが入力したいくつかのキーワードを組み合わせてることによってユーザの目的に合った情報を持つ Web ページを返すという直感的でわかりやすい検索機能、幅広いドメインで利用可能であることが挙げられる [1]。

しかし、ユーザは常に多数のキーワードを用いた検索クエリを正しく生成しているわけではない。Web 検索エンジン Excite<sup>\*3</sup>の、約 21 万ユーザにおけるクエリログ約 100 万件を分析した結果では、一つのクエリにつき平均 2.16 個のキーワードしか用いられておらず、ユーザは自分の情報要求を少ないキーワードで表現していることがわかる。すなわち、論理演算子を活用した複雑なクエリはあまりもちいられていないといえる。また、約 52 %のユーザが目的の検索結果を得るためにキーワード検索を繰り返していることが報告されている [2]。しかし、短いクエリが与えられた時には、語の持つ意味の多様性などにより、通常の検索では多様な内容をもつ Web ページが検索されてしまう問題がある [3]。また、AOL<sup>\*4</sup>の公開クエリログセットを分析した結果、Web 検索に入力される検索クエリのうち、約 28 %はそれ以前に入力されたクエリを修正したものであることが報告されている [4]。これらの結果から、多くのユーザは目的の情報を得るために複数のキーワードを用いて自分の情報要求を表現しながら検索を行い、目的の情報が得られなかった場合にキーワードの修正を行いながら検索を続けていることがわかる。しかし、多くのユーザは同時に多数のキーワードを指定せず、また一回の検索で目的の検索結果を得られていないことから、インターネット上の膨大な情報からユーザが自分自身の情報要求をキーワードで表現することは簡単ではないことがわかる。

これらの問題は、既存の Web 検索エンジンが提供する基本検索機能とキーワードを用いて表現するユーザの情報要求との乖離が大きく、個々の情報要求に合わせ、適切なキーワードに分解するのに要するユーザの負担が大きいたことが原因であると考えられる。

---

\*1 <https://www.google.co.jp/>

\*2 <https://www.yahoo.co.jp/>

\*3 <https://www.excite.co.jp/>

\*4 <https://www.aol.jp/>

### 2.1.2 高度な検索エンジン

既存 Web 検索エンジンの基本検索機能とユーザの情報要求との乖離の問題を解決するために、自然言語によるクエリ入力を受け付ける方法や、検索対象を特定のドメインに限定することで高度な検索を行う手法が研究されている。自然言語によるクエリ入力に関連した研究には、ユーザの自然言語による問い合わせに対して、システムが持つ知能を使って答える QA（質疑応答）システムがある [6][7][8]。これは、既存 Web 検索エンジンの様に情報要求をユーザ自身がキーワードの組み合わせとして指定することで検索を行うのではなく、文形式で入力し回答を返すシステムである。これによってユーザの負担を軽減することができる。現在では、IBM 社の Watoson や Apple 社の iOS に搭載された Siri, Microsoft 社の Windows に搭載された Cortana など、商用として広く利用されている。

質問応答システムが対象とする質問は、回答として得られる表現によって 2 つに分類される [5]。

- factoid 型：事実に対応するような、名称（人、組織、製品等の名前）や数に纏わる表現（金額、大きさ、日付等）など短い表現（名詞句）が回答となる。
- non-factoid 型：定義、理由、方法、関係など、名詞句よりも大きな単位（節、文、段落）が回答となる。複数箇所から得られた情報を統合・要約することもある。

このうち、商用システムなどで一般に用いられている手法は、factoid 型を対象としたものであり質問文に対して短く端的な回答を返すことを想定している。そのため、「理由」を尋ねる Why 型や「方法」を尋ねるの How 型のように回答が文章になる non-factoid 型質問への対応が困難であることが指摘されている [6]。

non-factoid 型質問を対象とした手法として、田村らは、Why 型質問に特化した自然言語検索を提案している [7]。この手法では、入力された Why 型質問文から、Web 検索エンジンに対する検索キーワードを作成するために質問文の言い換えを行う。入力された質問文に対して、疑問語（なぜ、どうして）の削除、文末表現（の、のですか。）の削除、助詞「は」の「が」への置換、文末への「のはなぜ」の追加、の 4 つの工程を行い検索エンジンへのクエリを作成する。検索結果のスニペットから情報を抽出し、回答を生成する。

三原らは、行動表現に着目し How 型質問に回答する質問応答を提案している [8]。この手法では、Web ページから「○○が△△したら、どうする？」のような、名詞の後に助詞、動詞と続き、さらに疑問詞が続く表現を質問文とし、検索エンジンのクエリとして用いる。検索結果の Web ページから「薬を塗る」のような「名詞 + 助詞 + 動詞」となる表現を抽出し、助詞を除いた表現（「薬：塗る」）を行動表現とし、回答として収集する。これらを事前に FAQ としてデータベース化することでユーザの質問に対し迅速な回答を実現している。自然言語によるクエリ入力はユーザのスキルを問わず利用可能であることが期待できるが、自然言語によるユーザの入力の自由度は高いため、それら全てに対応することは現状では困難である

既存 Web 検索エンジンが Web 上にある多様なページを対象としているのに対し、検索対象とするドメインを限定することで、効率的な検索の実現を目指した専門検索エンジンが提案されてい

る [9] [10][11][12]. ドメインを限定することで、そのドメインにおいて重要な情報を抽出して提示したり、キーワードのみをクエリとした Web 検索エンジンと比較して高度な検索機能を提供することができるといった利点がある.

WWW 上のコンピュータサイエンスに関する論文を対象とした専門検索エンジンが提案されている [9]. このシステムでは、コンピュータサイエンスに特化した強化学習により作成したロボット (クローラー) を用いることで、WWW 上から効率的に論文を収集している. 検索結果として論文のタイトルや著者名、要約を提示する他、論文の引用関係に基づきハイパーリンクを張ることで、関連論文へのアクセスを支援している.

亀井らは、WWW 上のソフトウェア開発に関する知見や情報に特化した検索エンジンを提案している [10]. この手法では、ソースコードや Tips のようなソフトウェア開発に役立つ知見・情報を含む Web ページを収集し、パッケージ名やクラス名、メソッド名やメトリクスに対して検索を行うことができる高度な検索機能を提供している.

Shakes らは、個人のホームページを対象とした検索エンジンを提案している [11]. この手法では、Web 検索エンジンを用いて検索を行うが、ユーザはキーワードだけでなく、個人の名前、組織、メールアドレス、国名などをクエリとして用いることができる. そして、検索エンジンから返された大量の検索結果に対し、組織の名前やメールアドレスなどから作成したドメインに特化したフィルタを用いて検索結果から不必要なページを削除することで、個人のホームページのみを検索結果として返す.

小久保らは多様なドメインに適応可能な専門検索エンジンを構築する手法として、「検索隠し味」を提案している [12]. 「検索隠し味」とは、決定木学習を用いて Web ページから抽出した、ドメインを限定するためのブール式である. 例えば、ユーザが豚肉を使ったレシピを検索したい場合、豚肉のキーワードの後に (AND 大さじ) の様な「検索隠し味」を追加することで、既存 Web 検索エンジンをレシピの検索に特化させることができる. この手法は Web 検索エンジンに「検索隠し味」を追加するだけで専門検索エンジンを構築することができるので、幅広いドメインで利用可能である.

検索対象をドメインに限定する専門検索エンジンでは、対象ドメインに関しては詳細かつ効率的な検索を行うことができる. しかし、構築に要するコストのため、多数のドメインに対して手法を応用することが困難であることが指摘されている [13].

### 2.1.3 コンテキスト検索エンジン

2.1.2 節で述べた専門検索エンジンは、ドメインを限定することで高度な検索機能を実現しているが、多様な人々が幅広い用途に利用可能であるという既存 Web 検索エンジンの長所が失われている. そこで、高間らは対象タスクを「動向に関する問い」に限定することで、ドメインを限定することなく、高度な検索機能を提供するコンテキスト検索エンジンを提案している [14]. ある商品の価格や売上、ある会社の業績状況等の時系列データをもととして、その変化を通時的にとらえつつ総合的にまとめ上げた情報は動向情報とよばれ [17], 世の中の様々な社会活動に深くかかわっている. 従って、動向情報を対象とすることで、幅広いドメインにおいて利用可能であることが期



待できる。コンテキスト検索エンジンでは、対象とする動向情報を、「コンテンツとしての動向情報」と「ユーザ活動による動向情報」の2種類に分類し、それぞれに関する時系列データを Web から収集してデータベースを構築する。「コンテンツとしての動向情報」とは、商品の価格や販売量に関するような、各企業や組織・団体によりコンテンツとして公開されている動向情報である。「ユーザ活動による動向情報」とは、WikipediaPageView 等の Web 上でのユーザ活動により発生する動向情報である。コンテキスト検索エンジンで収集している、「コンテンツとしての動向情報」の例を表1、「ユーザ活動による動向情報」の例を表2に示す。2018年1月の時点では、1,366,224アイテムが検索可能となっている。主要なデータであるWikipediaPageViewは、Wikipediaの各ページの1時間ごとのアクセス数をまとめたものが公開されており\*5、コンテキスト検索エンジンではこれを月単位で集計して利用している。Wikipediaに1日に1000回以上アクセスされたキーワードは、GoogleTrendsで得られた検索頻度との相関係数が0.72であり、WikipediaPageViewデータはWeb検索傾向と相関があるといわれている[18]。

表1 コンテンツとして動向情報の例 ([14] を参考に作成)

| 団体名           | 動向情報                   | 収集期間           |
|---------------|------------------------|----------------|
| 総務局統計局        | 人口，雇用者など               | 1953.1-2013.8  |
| 日本即席食品協会      | 生産数量（袋めんなど）            | 1958.1-2011.12 |
| 日本ミネラルウォーター協会 | ミネラルウォーター消費量           | 1986.1-2011.12 |
| NHK 放送文化研究所   | 内閣政党支持率                | 1998.4-2013.7  |
| 日本エネルギー経済研究所  | ガソリン卸価格                | 2000.7-2011.3  |
| 内閣府           | 景気ウォッチャー調査（百貨店，コンビニなど） | 2000.10-2011.3 |
| 農畜産業振興機構      | 価格（にんじん，キャベツなど）        | 2005.1-2012.8  |
|               | 消費者物価指数（にんじん）          | 2008.1-2010.11 |
| 自転車産業振興協会     | 自転車販売台数                | 2009.1-2011.12 |

\*5 <https://dumps.wikimedia.org/other/pagecounts-raw/>

表 2 ユーザ活動による動向情報の例 ([14] を参考に作成)

| 団体名         | 動向情報                    | 収集期間           |
|-------------|-------------------------|----------------|
| Yahoo!JAPAN | 急上昇ワードランキング (ポイント)      | 2007.11-2012.8 |
|             | ヒット数                    | 2003.1-2012.12 |
|             | ブログ記事数                  | 2005.1-2012.12 |
| Kizasi.jp   | きざしランキング (順位)           | 2006.4-2012.7  |
| ついっぶるトレンド   | HOT ワード (盛り上がり度, ツイート数) | 2010.3-2012.7  |
| Google      | Google Trends (検索数)     | 2004.1-2012.12 |
| Wikipedia   | WikipediaPageView       | 2008.1-2014.12 |

コンテキスト検索エンジンでは、「既存検索エンジンよりも動向に関して高度な検索が可能であること」、「複数の検索を組み合わせることで動向に関するユーザの多様な問いに答えられること」を設計方針とし、基本検索機能とクエリ構文を定めている。基本検索機能の設定には、既存 Web 検索エンジンを用いて二枚の画像から画像の撮影場所を特定してもらう課題を予備実験として行い [16]、ユーザの検索行動に対して考察を行うことで以下の 3 つの基本検索機能を実装している [15]。

Type1. 指定したアイテムに関する動向が特徴的変動を示した期間の検索

Type2. 指定した期間に特徴的変動を示したアイテム・動向の検索

Type3. 指定したアイテムに関する動向が特徴的変動を示した期間に同様の  
変動を示したアイテム・動向の検索

検索の際には、アイテム名、期間、特徴的変動を基本とし、検索の絞り込みオプションとして単位、リソースを組み合わせクエリを入力する。特徴的変動とは、時系列データが持つ時間変化のうち言語的に解釈可能なものであり、図 1 に示す 9 種類が指定可能である。

基本検索機能 Type1 ではアイテム名と特徴的変動をクエリとして指定することで、そのアイテムの動向で指定した特徴的変動が発生した期間が検索結果として返される。基本検索機能 Type2 では期間と特徴的変動をクエリとして指定することで、その期間内で指定した特徴的変動が発生したアイテムが検索結果として返される。基本検索機能 Type3 ではアイテム名と特徴的変動をクエリとして指定することで、そのアイテムで特徴的変動が発生している時期に同様の特徴的変動が発生するアイテムが検索結果として返される。全ての特徴的変動を検索対象とする事も可能である。検索結果のタイプは、期間の出力 (@period), アイテムの出力 (@item) が指定できる。基本検索機能の Type3 のみ、指定した特徴的変動が発生する動向と逆の動向を持つアイテムの出力 (@item opposite) も指定できる。コンテキスト検索エンジンのクエリ入力フォームを図 2 に示す。







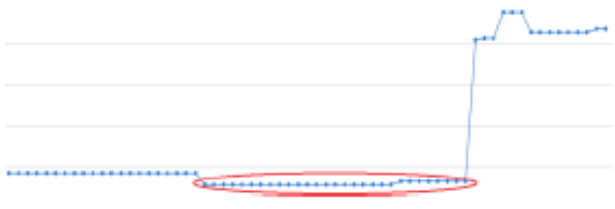


| 特徴的変動          | グラフ概形  | 抽出条件                             |
|----------------|--|----------------------------------|
| 最大値<br>(MAX)   |    | 各動向情報が最大の<br>期間                  |
| 最小値<br>(MIN)   |    | 各動向情報が最少の<br>期間                  |
| 急上昇<br>(SI)    |    | 3ヶ月以内に、変動幅<br>の20%以上の増加          |
| 急下降<br>(SD)    |   | 3ヶ月以内に、変動幅<br>の20%以上の減少          |
| ピーク<br>(PEAK)  |  | 変動幅の10%以上の<br>増加後、減少に転じた<br>山の頂点 |
| 底<br>(BOTTOM)  |  | 変動幅の10%以上の<br>減少後、増加に転じた<br>谷の頂点 |
| 安定<br>(NONE)   |  | 変動幅が月に1.5%を<br>超えない3ヶ月以上<br>の期間  |
| 基準値以上<br>(UP)  |  | 指定した基準値以上<br>の値をとる期間             |
| 基準値以下<br>(LOW) |  | 指定した基準値以下<br>の値をとる期間             |

図1 各特徴的変動のグラフ概形と抽出条件（文献 [21] を基に作成）

The image shows the input form of a Context Search Engine. It includes several dropdown menus and a search button. Red arrows and text labels point to specific parts of the form:

- 期間** (Period): Points to the date range selection area at the top, which includes **StartYear**, **StartMonth**, **EndYear**, and **EndMonth** dropdowns.
- アイテム名** (Item Name): Points to the **Item** dropdown menu, which has the example text "(e.g. 乾電池, にんじん, 自転車)".
- 特徴的変動** (Characteristic Variation): Points to the **Variation:ALL** dropdown menu.
- 検索結果のタイプ** (Search Result Type): Points to the **@Item** dropdown menu.
- 単位** (Unit): Points to the **Unit** dropdown menu.
- リソース名** (Resource Name): Points to the **All source** dropdown menu.

Other visible elements include a checkbox labeled "Only be used in variation UP/LOW" and a blue **Q Search** button.

図2 コンテキスト検索エンジンの入力フォーム

図3は、2013年1月から2013年6月までの期間で急上昇（SI）が発生したアイテムの検索結果上位3件である。通常のWeb検索エンジンと同様に、検索結果のランキング機能も提案されており、図3では動向情報の変動の激しさというコンテキスト検索エンジン特有の素性を用いたランキングスコアでソートされている[19]。検索結果の左からアイテム名（item）、単位（Unit）、リソース名（DBresource）、特徴的変動（VarType）、指定した特徴的変動が発生した期間（Period）、Sparklineを用いた動向情報の概要のグラフ（Value & Sparkline）、地域・対象（Target）、Googleボタン（Google）が提示されている。アイテム名をクリックすると、そのアイテムについての詳しい動向情報を示す時系列データのグラフを見ることができる。リソース名には動向情報の統計データ名（価格、WikipediaPageView等）が表示される。地域・対象では収集したアイテムが地域や対象で限定されているもの（「東京での」レギュラーガソリンの価格等）の場合、当該地域などが表示される。Sparklineを用いた動向情報の概要のグラフは、通常のWeb検索エンジンにおけるスニペットの役割を果たしており、検索結果の効率的な確認に効果があると報告されている[20]。Googleボタンをクリックすることで、そのアイテム名を検索クエリとしたGoogle検索へ飛ぶことができ、アイテムについての情報を得ることができる。また右端の星マークをクリックすることで、アイテムをブックマークすることができる。図4は進撃の巨人のWikipediaPageViewに関するグラフであり、詳細な値を確認することができる。

“ 2013/1-2013/6 SI @item” の検索結果 ( 668910 )




| # | Item  | Unit | DBsource           | VarType | Period            | Value & Sparkline   | Target | Google   |
|---|-------|------|--------------------|---------|-------------------|---|--------|----------|
| 1 | 進撃の巨人 | ビュー  | Wikipedia pageview | SI      | 2013-04 ~ 2013-05 | 479150.0<br>  |        | Google ☆ |
| 2 | 進撃の巨人 | ビュー  | Wikipedia pageview | SI      | 2013-02 ~ 2013-04 | 1142000.0<br> |        | Google ☆ |
| 3 | あまちゃん | ビュー  | Wikipedia pageview | SI      | 2013-02 ~ 2013-04 | 533590.0<br>  |        | Google ☆ |

図 3 コンテキスト検索エンジンの検索結果の例

### 進撃の巨人のWikipedia pageviewについてのグラフ

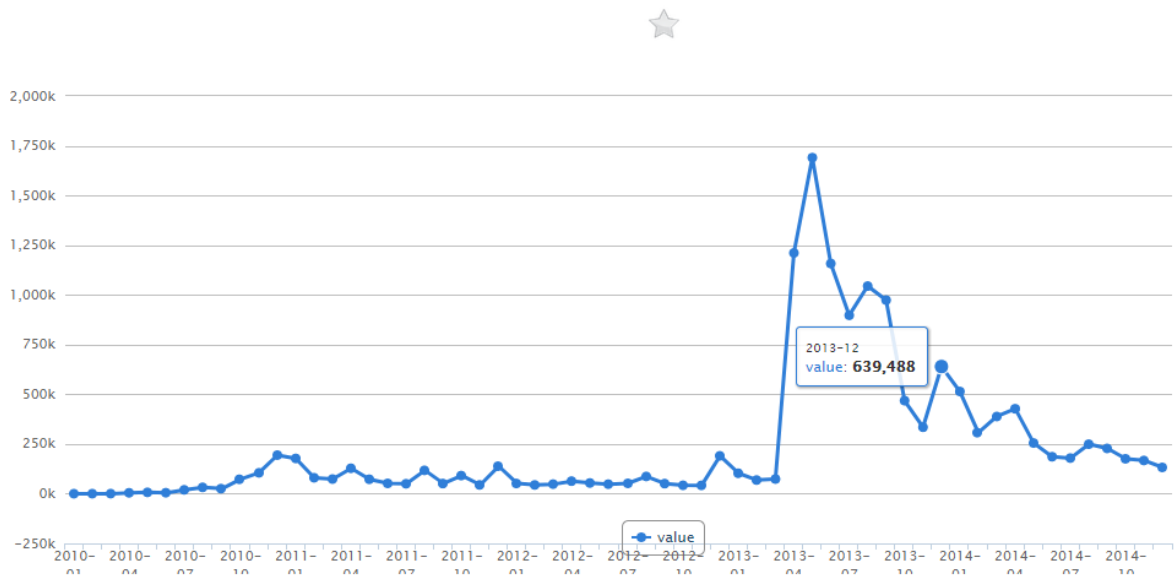


図 4 進撃の巨人の WikipediaPageView の動向を示すグラフ

## 2.2 Web 検索エンジンでの情報検索支援

Web ロボットと呼ばれるプログラムを用いて Web から収集した情報に対する全文検索を提供するロボット型の検索エンジンの台頭により、検索エンジンで検索可能な Web ページは膨大になった。その結果、検索結果として膨大なページが返されるようになり、その中から必要な情報が含まれているページを探す作業がユーザの負担となっている。そこで、従来の文献検索で用いられていた検索手法が導入されている。それに加えて、WWW の特性と情報検索手法とを関連付けた Web 検索エンジン特有の検索支援の開発が行われている。Web 検索エンジンで使われている情報検索手法についてまとめたものを表 3 に示す。検索エンジンに欠かすことのできない論理演算子は、Web 検索エンジンでも実装されており、以下の三つが利用可能である。

- AND：論理積を表す。(A AND B) で A と B の二つのキーワードが共に含まれている Web ページを返す。
- OR：論理和を表す。(A OR B) で A または B のどちらかのキーワードが含まれている Web ページを返す。
- NOT：論理差を表す。(A NOT B) で A というキーワードが含まれており、B というキーワードが含まれていない Web ページを返す。

ほとんどの Web 検索エンジンでは複数のキーワードを演算子なしで入力することで AND 検索を行う事ができる。しかし、多くのユーザは 2.1.1 節で述べた通り少ないキーワードしか用いず、論理演算を用いた検索結果の効率的な絞り込みは行っていない。特に日本語では複合語がフレーズ検索されるため、キーワードはさらに少なくなる傾向にあると言われている。日本語 Web サイトを主な検索対象とする Web 検索エンジン ODIN では、検索クエリの 7 割以上が一つのキーワードのみで構成されていることが報告されている [23]。

Web 検索エンジンで特に重要視されている情報検索支援としてランキングが挙げられる。従来の文献検索では、結果を全件見することを前提に検索結果を絞り込むことに主眼が置かれており、ランキングはオプションとして用意される程度であった [22]。しかし、Web 検索エンジンは膨大なページ数に対して全文検索を行っているので、有用な Web ページを検索結果の上位に表示するランキングの重要度が高まった。代表的なランキングの指標として、ページコンテンツを利用した指標 [26] やハイパーリンク構造を利用した指標 PageRank[24], HITS[25] などがある。ページコンテンツを利用した指標では、ページ中のキーワードの出現頻度、ページサイズの大きさ、キーワードの出現位置、被リンク数、スタッフによる評価などを総合したスコアを算出している。PageRank は、多くのページから引用されている Web ページは重要度が高く、また重要度が高いページから引用されている Web ページも重要度が高いという考えに基づいて、Web ページ間のハイパーリンクをランダムに選択しブラウジングするユーザ行動を想定してスコアを算出する。ハイパーリンクのみを対象として用いた静的なグラフ構造を用いているので、結果が入力クエリに依存しないため、事前にスコアを計算しておくことが可能である。HITS では、多くの Web ページへリンクし

ている Web ページを Hub, 多くの Web ページからリンクされている Web ページを Authority と定義し, この二つの役割に基づき Web ページの価値を評価していることが特徴である. 重要度の高い Authority へリンクしている Hub は重要度が高く, 重要度が高い Hub からリンクされている Authority は重要度が高いという再帰的な定義がなされており, これを用いて Web ページの重要度を決めている. 検索結果の Web ページ内でのハイパーリンクを用いた動的なグラフ構造を用いるので結果が入力クエリに依存し, 毎回の計算にコストがかかる.

表 3 検索エンジンで使われている手法 ([22] を参考に作成)

| 手法/技術          | 手法/技術の説明                         |
|----------------|----------------------------------|
| 論理演算子          | 複数の検索語関係を論理演算子を使い表現する            |
| フレーズ検索         | 複数語をそのままの形で検索する                  |
| 近接演算子          | 複数のキーワードの出現位置を指定する               |
| ランキング (順位付け出力) | 検索結果を適合度順に出力する                   |
| あいまい検索         | 入力文字列に厳密にマッチしない場合も検索結果に含める       |
| 自然言語処理         | 自然言語の形で入力された検索式から検索を行う           |
| フィールド指定        | Web ページ中の指定部分から検索を行う             |
| クエリ推薦 (検索式の拡張) | 検索式と関連が高いと思われる語を提示する             |
| 多言語検索          | 検索式の言語とは異なる kensak る言語のページを検索する  |
| マルチメディア検索      | 画像や音声ファイルを対象として検索を行う             |
| データベースの指定      | 分野, 有用性等から制限されたデータベースを対象とした検索を行う |

あいまい検索は, クエリとして入力された文字列を拡張し, 厳密にマッチしない場合でも検索結果として返す手法である. 大文字小文字や全角半角を区別せずに検索を行う手法や, 代替文字\* (アスタリスク) を用いることで一部不明なキーワードを含む検索を可能にする手法などが該当する.

フィールド指定は, タイトルや URL, アンカーテキストなどの Web ページタグ情報を用いて指定部分からの検索を行う手法である. 例えば, タイトルに「TMU」が含まれているページを検索したい場合, 検索コマンド「intitle」を用いた「intitle:TMU」をクエリとして検索を行うことで, 「TMU」という文字列がタイトルに含まれているページのみを検索結果として返すことができる.

しかし, 菅谷らが文系学部的女子大学生 1 年生から 4 年生の計 57 名に対して行った, コンピュータやインターネット利用に関するアンケート調査では, 上記の検索オプションを利用していると答えた学生は 0 人であった [27]. この結果から, Web 検索エンジンでは多様な検索オプションを用いた複雑な検索を行わなくとも, 検索結果に有用な Web ページを返すためのランキング機能がユーザに対しての重要な情報検索支援となっているといえる.

より詳細なクエリを入力してもらうことを目的とした支援機能として, クエリ推薦がある. クエ

リ推薦とは、ユーザの入力したクエリに対して別キーワードの追加や、クエリ内のキーワードを別のキーワードに置き換えることで、ユーザの情報要求により適したクエリを生成する機能である。Google<sup>\*6</sup>では、ユーザがクエリ入力フォームにキーワードを入力すると、検索回数や検索をした場所などから予測される、一緒に検索されやすいキーワードを候補キーワードとして表示するサジェスト機能や、検索結果の下に候補キーワードを表示する関連ワード機能がクエリ推薦として提供されている。Web 検索エンジン Yippy<sup>\*7</sup>では、検索結果のスニペットに対してクラスタリングを行うことで、検索クエリの特徴語を抽出し、ユーザに提示する機能が実装されており、これらの機能でユーザのクエリ生成に対して支援を行っている。

## 2.3 検索エンジンでの情報探索行動

Web 検索エンジンを用いて Web ページを検索する情報検索が主流となる前、情報検索の対象は雑誌論文等の文献が中心であり、主な利用者はその分野の研究者や図書館員であった [28]。当時の検索において、対象文献の主題に関する知識は不可欠であるが、それに加えて情報検索に関する専門的な能力が必要であり、エンドユーザに代わって検索を行うサーチャー（情報専門家）という職種が誕生した。やがて、オンライン蔵書検索システム（Online Public Access Catalog, OPAC）や CD-ROM による検索サービスが提供され、情報検索の知識は無いが専門的な情報ニーズを持っているエンドユーザが直接検索を行うようになった。さらに、情報検索の主流が Web ページに移ると、情報検索の利用者層は一変し、情報検索の知識も専門的な情報ニーズも持たないエンドユーザも Web 検索エンジンを用いて検索を行うようになった。

Jansen らは、従来の文献検索システムと OPAC, Web 検索に関する研究を比較し検索システムに見られる典型的な様相と Web の特殊性を論じている [29]。それらをまとめたものを表 4 で示す。

表 4 異なる情報検索システム研究の比較 ([29] を基に作成)

| 項目               | 従来の文献検索システム | OPAC   | Web   |
|------------------|-------------|--------|-------|
| 検索 1 回あたりのクエリ試行数 | 7～16 回      | 2～5 回  | 1～2 回 |
| クエリあたりのキーワード数    | 6～9 語       | 1～2 語  | 2 語   |
| 論理演算の使用率         | 37 %        | 1 %    | 8 %   |
| 検索エラー率           | 17 %        | 7～19 % | 10 %  |

この結果から、従来の文献検索システムは多くのキーワードと論理演算を用いて複雑な検索を行っているのに対し、OPAC や Web では、簡単な検索しか行われていないことがわかる。また、従来の文献検索システム、OPAC, Web の順にクエリ試行数が減少しており、ユーザの情報要求自体も簡単になっていることがわかる。

<sup>\*6</sup> <https://www.google.co.jp/>

<sup>\*7</sup> <https://yippy.com/>



種市らは、大学生を対象とした Web 検索エンジンでの検索実験を行いその結果から、典型的な情報探索行動のパターンとして以下の点を挙げている [30].

- 2 語程度のキーワードで検索を行い、結果は最初の 1, 2 ページのみを確認する.
- 検索結果のヒット件数はほとんど確認せず、結果の上位に表示される情報を選ぶ.
- 結果の選別は視覚的要素と経験的要素を基に行われており、コンテンツの質的評価は欠落している.

このような状況を踏まえ、専門的なスキルを持たないエンドユーザが精度の高い検索を行い、質の高い情報を獲得するためにはユーザ自身のスキル向上が必要との指摘がなされている [31].

情報検索に関する専門的なスキルを駆使して、エンドユーザの代わりに検索を行うサーチャーは、情報検索に関する知識以外にも、一般的なデータベースに関する知識、個々の文献データベースに関する知識とその検索技能、エンドユーザの求める情報の主題に関する知識、結果として得られた情報を活用する知識・技能などが必要であるといわれている [32]. サーチャーはエンドユーザの情報要求に対して図 5 のステップに沿って情報検索を行う.

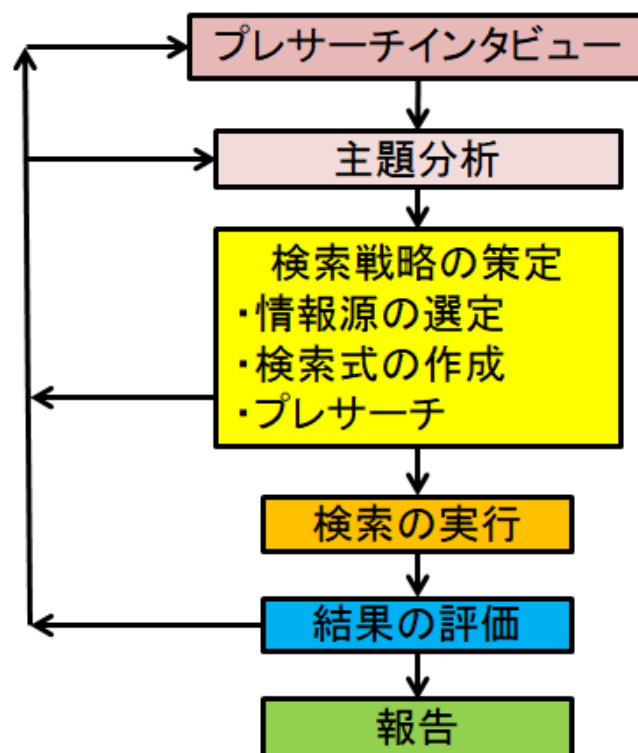


図 5 サーチャーの情報検索のプロセス（文献 [33] を基に作成）

サーチャーはまずエンドユーザの意図や情報要求を理解し、結果をどのように活用するかなどを確認するプレサーチインタビューを行う。次に、エンドユーザの意図する情報を適切に検索するために、主題となる要素を適切な言葉や分類などで表現するための主題分析を行う。検索戦略の策定

では、適切な情報源の選定、検索式の作成などを具体的に決めていく。この時に、情報要求の目的に適合した情報源の選択などの「正確性」、目的の情報を漏れなく取り出す「網羅性」、依頼期日に合わせた検索スケジュールを策定する「迅速性」の3つが要点として挙げられる。この際に適切な情報源を選択するためのプレサーチを行うこともある。その後、策定した検索戦略に沿った検索を行う。検索結果を評価する指標として以下の2つがある。

- 適合率： 検索結果の文献数に対する、適合した文献数の比率
- 再現率： データベース内の全適合文献数に対する、検索結果の適合文献数の比率

サーチャーは常に高い適合率と再現率を目指す、目的に応じてどちらを優先すべきかという視点に立ち戦略を立てる。例えば、医療や特許の検索において再現率は特に重要である。医療検索では、副作用情報を一つでも漏れてしまうと重大な医療ミスにつながる [34]。特許検索では、重要な出願が1件でも漏れていたら特許侵害に繋がってしまう [35]。このような場合、ノイズを多くしてでも再現率の高い検索結果が得られるような、網羅的な検索戦略を立てる必要がある。これらの工程を経て、得られた結果がエンドユーザの情報要求に満たしているかを評価し、適切でないと判断された場合には再びプレサーチインタビューや主題分析から検索をやり直す。そして、情報要求を満たしていると判断された結果をまとめエンドユーザに返すまでがサーチャーの情報検索の一連の流れになる。以上の様に、サーチャーはエンドユーザによる Web 検索での検索と異なり、複雑な検索式を作成することで目的に合わせた検索結果を作成し、その結果を網羅的に確認するという高度な情報探索行動を行っていることがわかる。

### 3 コンテキスト検索エンジンでの効率的な絞り込み支援

本章では、コンテキスト検索エンジンにおける効率的な絞り込み支援を実現するアプローチを提案する。2.1.3 項で説明したコンテキスト検索エンジンは、対象タスクを「動向に関する問い」に限定することで、ドメインを限定することなく高度な検索機能を提供している。しかし、コンテキスト検索エンジンでは時系列データを検索対象としており、それに対応した独自のクエリ構文を入力する必要があるため、クエリ生成に要するユーザの負荷が課題となっている。また、データベースの拡充によって 100 万以上のアイテムに対して検索を行うことができるようになったが、それに伴い検索結果として返されるアイテムも増加しており、その効率的な絞り込みも課題となっている。

そこで本論文では、コンテキスト検索エンジンの効率的な絞り込み支援として、論理演算機能の導入を提案する。また、コンテキスト検索エンジンの効果的な活用を支援するために、コンテキスト検索エンジンのガイドラインを作成することで、ユーザに対して適切な知識を提供するとともに、効率的な検索行動を提示する。本章では、3.1 節でコンテキスト検索エンジンへの論理演算機能の導入、3.2 節でコンテキスト検索エンジンのガイドラインの作成、3.3 節でコンテキスト検索エンジンでの効率的な検索行動支援を提案する。

### 3.1 コンテキスト検索エンジンへの論理演算機能の導入

本節では、コンテキスト検索エンジンへの論理演算機能の導入を提案する。論理演算機能は従来の検索エンジンで一般的に用いられている検索結果の絞り込み手法の一つである。しかし、Web 検索エンジンと従来の文献検索システムではクエリ生成で異なる点がある。Web 検索エンジンでは、クエリ入力フォームにキーワードを入力しキーワードが含まれている Web ページを返すシステムだが、探索状態が保持されず結果を積み重ねて構造化することができないという特徴があり、1 回ごとにクエリを生成し検索結果をみてクエリを修正しなければならない。これに対し、従来の文献検索システムでは対話型のシステムを取り入れており、探索状態を保持する機能を持っている。探索状態の保持とは、行われた検索のクエリと検索結果を記録していくことであり、それらを確認した結果に基づき次のクエリを生成したり、検索結果同士の論理演算を行うことができる。対話型検索システムの例として、生命科学に関する書誌情報データベース Medline の検索システム「Pubmed」[36] の使用例を図 6 で示す。

The image shows the PubMed search interface. At the top is the 'Builder' section with three rows of search criteria. The first row has 'All Fields' and '(o157) AND prevention'. The second row has 'AND' selected from a dropdown, 'All Fields', and '(winter) OR summer'. The third row has 'AND' selected, 'All Fields', and an empty field. To the right of each row is a 'Show index list' link. Below the builder is a 'Search' button and a link to 'Add to history'. Below this is a red heading 'クエリ入力フォーム'. Underneath is the 'History' section, which is a table with columns: Search, Add to builder, Query, Items found, and Time. The table lists 9 previous searches, each with an 'Add' link. At the top right of the history section are links for 'Download history' and 'Clear history'.

| Search | Add to builder      | Query  | Items found             | Time     |
|--------|---------------------|--|-------------------------|----------|
| #9     | <a href="#">Add</a> | Search (((o157) AND prevention)) AND ((summer) AND winter) | <a href="#">8</a>       | 11:32:09 |
| #8     | <a href="#">Add</a> | Search (((o157) AND prevention)) AND ((winter) OR summer)  | <a href="#">44</a>      | 11:31:39 |
| #7     | <a href="#">Add</a> | Search (winter) OR summer                                  | <a href="#">207806</a>  | 11:31:30 |
| #6     | <a href="#">Add</a> | Search (summer) AND winter                                 | <a href="#">17958</a>   | 11:30:57 |
| #5     | <a href="#">Add</a> | Search winter  | <a href="#">120031</a>  | 11:30:47 |
| #4     | <a href="#">Add</a> | Search summer  | <a href="#">105733</a>  | 11:30:42 |
| #3     | <a href="#">Add</a> | Search (o157) AND prevention                               | <a href="#">1495</a>    | 11:30:30 |
| #2     | <a href="#">Add</a> | Search prevention  | <a href="#">1620116</a> | 11:30:19 |
| #1     | <a href="#">Add</a> | Search o157  | <a href="#">10483</a>   | 11:30:09 |

### 探索情報を保持している部分

図 6 Pubmed の利用例

図 6 では、o157 の予防と季節についての検索を行っている。上の Builder の部分がクエリフォームである。クエリ入力フォームの一段目には o157 と prevention を AND 演算子で組み合わせたクエリが入力されており、その左側のプルダウンボックスで検索対象フィールドを指定することができる。二段目には winter と summer を OR 演算子で組み合わせたクエリが入力されており、一番左に論理演算子を選択するプルダウンボックスが表示されている。ここで、一段目のクエリと二段目のクエリをどの論理演算子でつなぐか選択することができる。「Search」ボタンをクリックすることで検索が実行され、「Add to history」をクリックすることで、クエリと検索結果件数を下の History に追加することができる。この History の部分が探索状態を保持している部分になっている。History 内では各クエリに対して番号が振られており、Query 列には検索クエリ、items found 列には検索結果件数が表示されている。Add to builder をクリックすることで、そのクエリを上の入力フォームに追加することができる。図 6 の例では、# 3 のクエリと # 6 のクエリをクエリ入力フォームに追加している。この 2 つのクエリを AND でつないで検索を行った結果が # 8 であり、44 件の文献がヒットしている。このように、対話型のシステムでは検索は段階的に行われ、検索結果間の論理演算の方が重要になっている。

Web 検索エンジンでは、キーワードに対して論理演算を行うのに対し、文献検索システムでは検索結果に対して論理演算を行う手法を採用している。Web 検索エンジンの手法をコンテキスト検索エンジンで適用する場合、独自に定義された構文に従ったクエリを一度に複数個生成しなければならず、ユーザのクエリ生成に対する負担が大きくなってしまふことが考えられる。複数の特徴的変動などを同時に指定するよりも、入力クエリに対する検索結果を確認したうえで次の検索クエリを生成することができる文献検索システムの手法が、着目すべき期間や変動を個別に検討できるため、コンテキスト検索エンジンには適していると考えられる。

そこで本論文では、文献検索システムと同様の論理演算機能を導入することで、クエリ生成の負担を増加させることなく効率的な検索結果の絞り込みを実現する。図 7 では、論理演算機能を実装したコンテキスト検索エンジンのクエリ入力フォームを示す。

図 7 論理演算機能を実装したコンテキスト検索エンジンのクエリ入力フォーム

検索結果のタイプを指定するフォームの右側に、論理演算子を選択するプルダウンボックスを追加している。選択できる論理演算子は以下の4種類である。

- NONE：探索情報をリセットし検索を行う。
- AND：探索情報として保持している検索結果と現在のクエリの結果、どちらにも含まれる検索結果を返す
- OR：探索情報として保持している検索結果と現在のクエリの結果、少なくともどちらか一方の検索結果に含まれる検索結果を返す。
- NOT：探索情報として保持している検索結果から現在のクエリの検索結果を除いた結果を返す。

これらを選択することで、検索結果の論理演算を行うことができる。提案手法では探索情報としてクエリとその検索結果件数を保持している。論理演算子を選択して実行した検索結果も、新たな探索情報として保持される。論理演算した結果の例を図8で示す。



図8 論理演算をした検索結果例

1 つめのクエリでは、2011 年 1 月～2011 年 3 月の期間に特徴的変動 PEAK が発生しているアイテムを検索している。クエリの右側には、その時点での検索結果件数が表示されており、このクエリの検索結果が 116,676 件であることがわかる。2 つめのクエリでは、2012 年 1 月～2012 年 3 月の期間に特徴的変動 PEAK が発生しているアイテムを検索している。このクエリでは、論理演算子 AND が選択されており、1 つめの検索結果にも含まれているアイテム 16,782 件が検索結果として返されている。3 つめのクエリでは、2013 年 1 月～2013 年 3 月の期間に特徴的変動 PEAK が発生しているアイテムを検索している。このクエリでは、論理演算子 NOT が選択されており、2 つめの検索結果からこのクエリの検索結果に含まれているアイテムを除いたアイテム 14,136 件が検索結果として返されている。この検索では、2011 年 1 月～3 月と 2012 年 1 月～3 月に PEAK が発生しており、2013 年 1 月～3 月わかるに PEAK が発生していないアイテムが検索結果として返されており、図 8 の検索結果の Sparkline を確認すると周期的な変動ではなく、当該時期に PEAK が発生しているアイテムが返されていることがわかる。この結果を確認すると論理演算を利用することで検索結果を絞り込めていることがわかる。

提案手法ではユーザが現在の検索結果に追加して次の検索を行うことを想定し、クエリ入力複雑さを軽減するために選言標準形での検索を採用する。選言標準形とは、論理式の標準化の一つであり、論理積 (AND) の論理和 (OR) の形式で論理式を表す。選言標準形を用いることで、各演算子の優先順序を指定しなくても解釈のあいまいさを回避することができる。本システムでは、例えば「A and B or C」のような検索を行った場合、これを選言標準形とみなし「(A and B) or C」と解釈する。

論理演算機能の導入に伴い、クエリの削除や論理演算子の変更を行うための編集ページを作成する。図 8 では、検索結果の右上に「クエリを修正する」というリンクがあり、これをクリックすることで編集ページに飛ぶことができる。図 8 の状態に対する編集ページを図 9 に示す。

The image shows a web interface for editing queries. It contains three query entries, each with a '削除' (Delete) button and a text field. Below each query is a row of radio buttons for logical operators: AND, OR, and NOT. The first query is '2011/1-2011/3 PEAK @item' with 'AND' selected. The second query is '2012/1-2012/3 PEAK @item' with 'NOT' selected. The third query is '2013/1-2013/3 PEAK @item'. At the bottom is a large '実行' (Execute) button.

|    |                          |   |
|----|--------------------------|---|
| 削除 | 2011/1-2011/3 PEAK @item | <input checked="" type="radio"/> AND <input type="radio"/> OR <input type="radio"/> NOT |
| 削除 | 2012/1-2012/3 PEAK @item | <input type="radio"/> AND <input type="radio"/> OR <input checked="" type="radio"/> NOT |
| 削除 | 2013/1-2013/3 PEAK @item |   |
| 実行 |                          |   |

図 9 図 8 の状態に対する編集ページ

図 9 では、各クエリの左側に削除ボタンがあり、これをクリックすることで該当するクエリを削除することができる。クエリ間の論理演算子をラジオボタンで変更し、一番下の実行ボタンをクリックすることで検索を行うことができる。例として、図 9 の一番上のクエリを削除し、残りの 2 つのクエリを OR 演算子で結合した検索結果を図 10 で示す。編集ページを利用して検索を行うと、図 10 のように編集前の全クエリについての探索情報は削除され、新たなクエリが検索結果件数と共に探索情報として保持される。

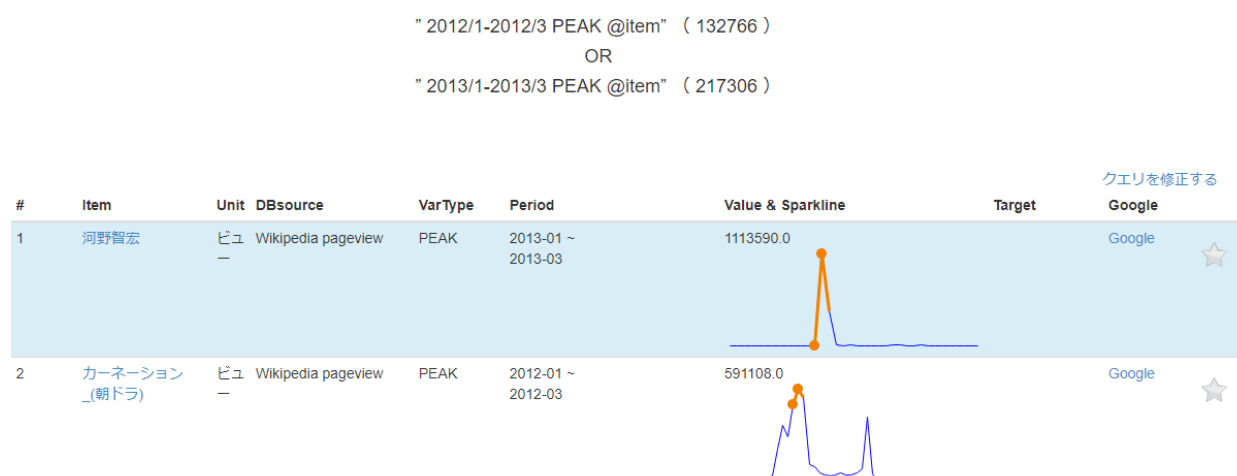


図 10 図 9 の編集後の検索結果

4.1 節の比較実験において、編集ページを用いたクエリの編集を行った実験協力者が少なかった。クエリの編集機能があまり利用されなかった要因の一つとして、クエリ編集ページに移動しなければクエリの編集が行えなかったことが挙げられる。そこで、図 11 に示す、検索結果画面にクエリ編集機能を導入する。図 11 では、検索クエリの下に「クエリの修正」ボタンが追加されており、これをクリックすることでクエリの編集フォームが表示される。



クエリ編集フォーム

[クエリを修正する](#)

図 11 検索結果でのクエリ編集機能

### 3.2 コンテキスト検索エンジンのガイドラインの作成

本節では、コンテキスト検索エンジンのガイドラインの作成について述べる。2.1.3 節で述べた通り、コンテキスト検索エンジンは独自に定義された構文に従ってクエリを生成しなければならない。システムについて詳しい知識を持たないユーザにとって、適切なクエリを生成することは負担が大きい。また、従来の Web 検索エンジンはキーワードを検索クエリとし、キーワードが含まれている Web ページを検索結果として返すのに対し、コンテキスト検索エンジンはアイテム名や期間、特徴的変動などを検索クエリとし、動向情報を検索結果として返しており、検索を行うには従来の情報検索とは異なる知識が必要である。

そこで、コンテキスト検索エンジンのガイドラインを作成することでユーザに対して基本的な利用方法についての適切な知識を提供する。また、3.1 節で追加した論理演算機能についてのガイドラインも作成することで、論理演算機能の基本的な利用方法も提供する。

コンテキスト検索エンジンのガイドラインの内容としては、特徴的変動と 3 種類の基本検索機能、アイテム名の部分一致検索機能について説明を行う。特徴的変動の説明では、各特徴的変動のグラフ概要と抽出条件、期間指定を行う時の注意点について示す。3 種類の基本検索機能については、使用例を載せており、実際のシステムにクエリを入力、検索を行ってもらうことでシステムの理解を促す。アイテム名の部分一致検索機能は、本研究において新たに実装した機能であり、この機能を用いてアイテム名での検索を完全一致検索か部分一致検索か指定することができる。図 12 は、部分一致検索機能を導入したクエリ入力フォームである。アイテム名入力フォームの左側に完全一致検索か部分一致検索か選択できるプルダウンボックスが追加されている。

The image shows a search interface with several input fields and a search button. At the top, there are date range selectors: 'StartYear' and 'StartMonth' followed by a tilde '~' and 'EndYear' and 'EndMonth'. Below these is a search type dropdown menu that is open, showing two options: '完全一致' (Exact Match) and '部分一致' (Partial Match). The '部分一致' option is currently selected and highlighted in blue. To the right of the dropdown is a text input field containing 'Item (e.g. 乾電池, にんじん, 自転車)' and a 'Unit' label. Further right is another dropdown labeled 'All source'. Below the search type dropdown is a text input field with the placeholder 'input AVG or any number'. Below that is a dropdown labeled '@Item' and another labeled '@AND'. At the bottom center is a blue button with a magnifying glass icon and the text 'Search'.

図 12 アイテム名の部分一致検索機能

論理演算機能のガイドラインでは、論理演算子「NONE」、「AND」、「OR」、「NOT」と検索結果ページ及び編集ページにおけるクエリの編集方法を説明する。実際の使用例とそれに対する操作方法を載せており、実際にシステムを動かすことで論理演算機能の理解を促す

4.2 節の予備実験で用いたガイドラインを付録 A で示す。

### 3.3 コンテキスト検索エンジンでの効率的な検索行動支援

本節では、コンテキスト検索エンジンでの効率的な検索行動支援を提案する。論理演算機能が導入される以前のコンテキスト検索エンジンでは、1つのクエリのみを用いた検索しか行うことができなかったが論理演算機能を導入することで、複数のクエリを組み合わせた複雑な検索が可能になる。しかし、2.3節で述べた通り、既存のWeb検索エンジンでは論理演算子を用いた複雑な検索は行われず、一つのキーワードのみを用いた単純な検索を行うことが多いことが知られており、この問題点は、コンテキスト検索エンジンでも当てはまることが想定される。

そこで、ユーザの検索行動を収集し、分析することでコンテキスト検索エンジンを利用した効率的な検索行動を明らかにし、ユーザに提示することで、コンテキスト検索エンジンの効果的な活用を支援する、

ユーザの様々な検索行動を収集するため、ユーザの興味によって検索タスクを決定する、従来のWeb検索エンジンの利用に近いタスクを想定した予備実験を行い、クリックログやクエリログなどから検索行動を収集する。

コンテキスト検索エンジンを効率的に利用してもらうために、ユーザに対して以下の3つの知識を提供する。

1. コンテキスト検索エンジンの検索システムについての知識
2. 検索結果として返される「動向情報」に対する知識
3. 情報検索に対する知識

4.2節に示す予備実験で収集したユーザのクエリログを確認し、適切な検索結果を得られていなかったクエリを分析した結果、以下の4点についてコンテキスト検索エンジンのクエリ生成に対して誤った認識をしている可能性を発見した。

- 検索結果のタイプ選択について
- アイテム名検索での検索結果について
- 特徴的変動「MAX, MIN」と論理演算について
- アイテム名の部分一致検索の利用方法について

そこで、これらに関する正しい知識を「コンテキスト検索エンジンの検索システムについての知識」として提供する。これによりユーザのコンテキスト検索エンジンに対する理解を促し、検索結果が0件や膨大な件数になることを防ぐ効果がある考えられる。

予備実験で発見された情報をユーザごとにまとめ、分析を行った結果、従来のWeb検索エンジンの検索結果として返されるWebページに比べて、コンテキスト検索エンジンの検索結果として返される動向情報はユーザにとってなじみが薄く、返された検索結果から有益な情報を取り出すことが困難である事がわかった。そこで、検索結果の効率的な探索に有効と考える以下の2点を「動向情報に対する知識」として提供する。

- 検索結果中の有益な情報の判断
- 検索結果中のノイズの判断

これらの知識を提供することで、動向情報に対する適切な知識をユーザに提供し、検索結果からの情報発見を支援する。

上記の2種類の知識は、コンテキスト検索エンジンに依存する知識であった。これら以外にサーチャーが従来の文献検索システムで行ってきた主題分析や検索結果の評価などの検索行動も、検索初心者にとっては有効と考える。そこで、以下の4点についてコンテキスト検索エンジンを前提とした行動に変換したものを、「情報検索に対する知識」として提供する。

- 情報要求を適切に表現するための主題分析
- 検索結果件数を利用した検索式の作成
- 検索結果の確認方法
- 網羅的な検索結果の確認

これらの知識を提供することで、ユーザの情報検索に対する能力の向上が見込める。

考察から得られた効率的な検索行動に対する知識を実験協力者にフィードバックして再度実験を行ってもらうことで、コンテキスト検索エンジンの効率的な活用支援に効果があるかを検証する。

本節で作成した効率的な検索行動に対する知識を付録Bで示す。

## 4 評価実験

本章では，提案する論理演算機能の有効性を検証するため 2 つの評価実験を行った結果を示す．

4.1 節では，論理演算機能を導入したコンテキスト検索エンジンと従来のコンテキスト検索エンジンとの比較実験の結果を示し，論理演算機能の有効性について考察する．4.2 節では，ログ収集のために行った予備実験の結果を示し，効率的な検索行動に対して考察を行う．4.3 節では，4.2 節の考察から得られた効率的な検索行動に対する知識をユーザに提示したうえで行った評価実験の結果を示し，ユーザに対する効率的な検索策行動支援の有効性について考察する．

### 4.1 論理演算機能を導入したコンテキスト検索エンジンに関する比較実験

#### 4.1.1 実験概要

工学系大学生・大学院生 8 名に対し，以下の 4 つの検索タスクをまとめた実験タスクを 2 回行ってもらった．

- a). 年単位で周期的なアイテムの発見
- b). 複雑な動向を示すアイテムに関連するアイテムの発見
- c). 時期の異なる同一イベントに関連するアイテムの発見
- d). 周期的ではないアイテムの発見

各実験タスクで回答してもらう内容を以下に示す．

#### 1. 実験タスク 1

- (a) 毎年 5 月に関心が高まっているアイテム (3 つ) およびその理由
- (b) 「インフルエンザ」の関心が高まっているときに関心が高まっているアイテム (3 つ) およびその理由
- (c) 2008 年、2012 年の夏季オリンピックに 2 大会連続で出場している選手 3 人（できれば，メダルを獲得している選手）
- (d) 2012 年に周期的にではなく関心が高まっているアイテム (3 つ) およびその理由

#### 2. 実験タスク 2

- (a) 毎年 6 月に関心が高まっているアイテム (3 つ) およびその理由
- (b) 「台風」の関心が高まっているときに関心が高まっているアイテム (3 つ) およびその理由
- (c) 2010 年、2014 年の冬季オリンピックに 2 大会連続で出場している選手 3 人（できれば，メダルを獲得している選手）
- (d) 2013 年に周期的にではなく関心が高まっているアイテム (3 つ) およびその理由

タスクの詳細については次項で説明する．各実験協力者には，3.2 節で作成したコンテキスト検

検索エンジンのガイドラインを読んでもらった後、この実験タスクを論理演算機能を導入したコンテキスト検索エンジンと従来のコンテキスト検索エンジンの両方で行ってもらう。各検索タスクの制限時間は 8 分に設定した。実験タスクと論理演算機能の有無の組合せ、およびタスクの実行順序は表 5 に示す通り、実験協力者によって異なるように設定した。実験後には、論理演算機能に関するアンケートに回答してもらった。従来の Web 検索エンジンの利用は、コンテキスト検索エンジンで発見したアイテムに関する情報を検索する場合のみ許可した。

コンテキスト検索エンジンのクエリログを確認すると、特徴的変動の「UP」と「LOW」の利用頻度が著しく低く、またこの 2 つの特徴的変動のみ他の特徴的変動と異なる方法で実装されている。そこで本論文では「UP」「LOW」の論理演算機能は実装しなかった。

表 5 実験タスク・論理演算機能の有無の組合せ、およびタスクの実行順序

| 協力者 ID | 1 回目           | 2 回目           |
|--------|----------------|----------------|
| 1,2    | 実験タスク 1・論理演算なし | 実験タスク 2・論理演算あり |
| 3,4    | 実験タスク 1・論理演算あり | 実験タスク 2・論理演算なし |
| 5,6    | 実験タスク 2・論理演算なし | 実験タスク 1・論理演算あり |
| 7,8    | 実験タスク 2・論理演算あり | 実験タスク 1・論理演算なし |

#### 4.1.2 検索タスクの説明

この項では、4 つの検索タスクについて説明を行う。

検索タスク (a) は「年単位で周期的なアイテムの発見」である。野菜の価格や自転車の販売量、Google 検索数など周期性を持つ動向情報がある一方で、乾電池の価格や内閣支持率など周期性の見られない動向情報がある。周期性を持つ動向情報の特徴は、季節や気候、クリスマスやバレンタインデーといった定期的な行事のような、周期性を持って発生する要因に影響を受けている点である。これらの要因について関心がある場合、周期性を持つ動向情報は価値ある情報であると考えられている [19]。従来のコンテキスト検索エンジンでは複雑なクエリを生成できないので、周期性のあるアイテムのみを検索することは困難であったが、論理演算子を用いることでそのようなアイテムを検索することが可能になる。周期性のあるアイテムについて検索を行った検索結果の例を図 13 で示す。

図 13 では、毎年 2 月に PEAK が発生しているアイテムを検索している。この例では、論理演算子 AND を用いることで 2013 年 1 月～2013 年 3 月、2012 年 1 月～2012 年 3 月、2011 年 1 月～2011 年 3 月の 3 つの期間で特徴的変動 PEAK が発生しているアイテムを検索している。検索結果のアイテム名を確認すると、「バレンタインデー」、「恵方巻」、「節分」と 2 月に関連のあるアイテムが上位に来ていることがわかる。また、各アイテムの Sparkline を確認すると一定間隔毎に値が上昇・下降しており周期性があることがわかる。このように、検索クエリの特徴的変動と期間の月を変えずに、期間の年だけを変えた検索クエリを論理演算子 AND を用いて組み合わせること

で、年単位で周期性のあるアイテムを発見することができる。そこで、実験タスク1では毎年5月に関心が高まっているアイテムを3つ、実験タスク2では毎年6月に関心が高まっているアイテムを3つ発見してもらい、それぞれのアイテムについての理由を回答してもらった。

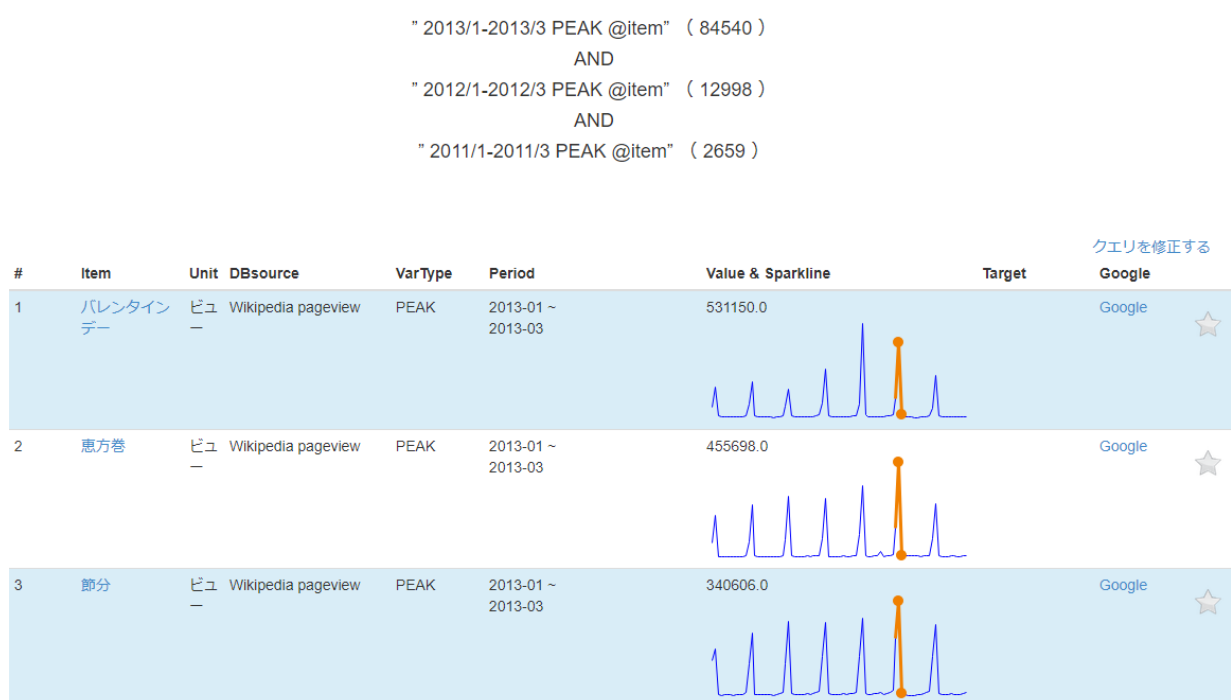


図 13 毎年2月にPEAKが発生しているアイテムの検索

検索タスク (b) は「複雑な動向を示すアイテムに関連するアイテムの発見」である。従来のコンテキスト検索エンジンでは、時間的変化の観点から関係のあるアイテムを発見するタスクで有効性が示されている。しかし、データ量の増大によって多数の検索結果が返されるようになり、時間的変化の観点から関係のあるアイテムであってもその発見が困難になっている。この問題に対し論理演算機能を導入したコンテキスト検索エンジンでは、複数のクエリを組み合わせることで複雑な動向情報を表現することが可能になる。アイテム「台風」に関連するアイテムの発見を例に説明を行う。まず、台風の動向情報を確認することでどの時期にどの特徴的変動が発生したかを理解する。台風の WikipediaPageView に関するグラフを図 14 に示す。

図 14 で丸で囲まれている箇所は特徴的変動 PEAK が発生している期間である。赤い丸の期間は9月～10月であり、これは台風が日本に接近することが多い期間と重なっている。この結果から、台風が日本に接近する期間に周期的に人々の関心が高まっていることがわかる。青い丸は9月～10月以外の期間に発生している PEAK である。これらの非周期的に PEAK が発生している期間では、台風が時期外れではあるが日本に接近している。例えば、2012 年 6 月に PEAK が発生し

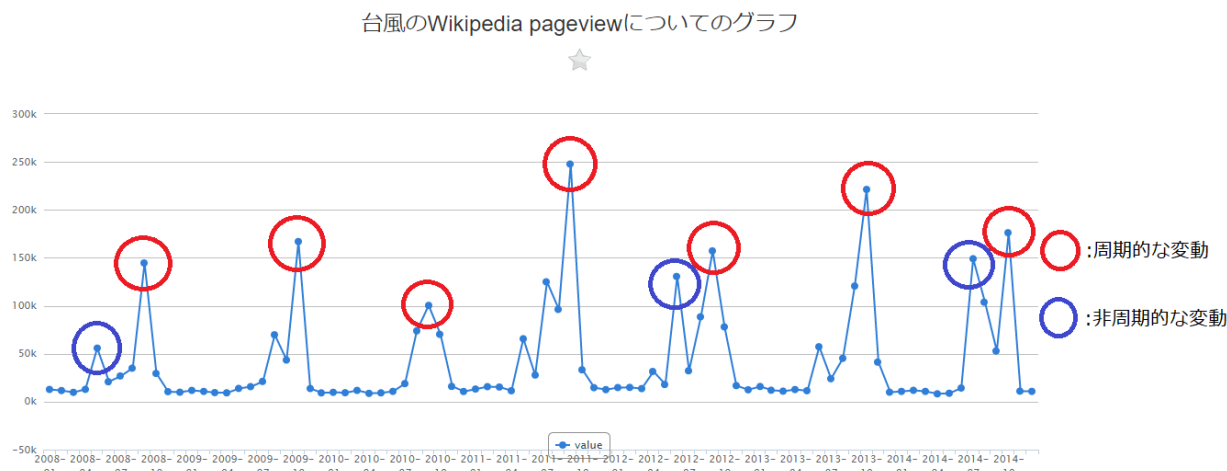


図 14 台風の WikipediaPageView の動向を示すグラフ

ているが、この時期に季節外れの台風 4 号が本州に上陸している\*<sup>8</sup>。このように周期的な変動と非周期的な変動が組み合わされた複雑な動向情報を持つアイテムも存在する。台風の複雑な動向情報を表現したクエリで検索した結果を図 15 で示す。

図 15 では、4 つのクエリを組み合わせて検索を行っている。赤枠で囲われた 2 つのクエリはどちらも 9 月に PEAK が発生しているアイテムを検索しており、台風の周期的な変動を表現している。青枠で囲われた 2 つのクエリは 2012 年 6 月と 2014 年 7 月に PEAK が発生しているアイテムを検索しており、台風の非周期的な変動を表現している。このように、周期的な変動・非周期的な変動を組み合わせたクエリで台風の動向情報を表現している。検索結果を確認すると、「台風」、「熱帯低気圧」、「温帯低気圧」のような「台風」に関係のあるアイテムが上位に来ていることがわかる。この検索結果の上位 20 件のアイテム名を表 6 で示す。

表 6 図 15 の検索結果上位 20 件のアイテム名

|           |            |                |                |
|-----------|------------|----------------|----------------|
| 台風        | 熱帯低気圧      | 温帯低気圧          | えなりかずき         |
| 台風の日      | 洞爺丸台風      | 気象警報           | 狩野川台風          |
| 沖永良部台風    | 伊東深水       | カスリーン台風        | ジェーン台風         |
| 宮古島台風     | キャロル・ベイカー  | 平成 10 年台風第 7 号 | 大植英次           |
| 第 3 宮古島台風 | 順天中学校・高等学校 | ウィリー・ウィリー      | 平成 8 年台風第 12 号 |

表 6 で赤字のアイテムは台風に関連したものであり、検索結果上位 20 件のうち 15 件を占めている。この結果から、台風の動向情報を適切にクエリとして表現することで、台風と関連のあるアイテムを発見可能であることがわかる。実験タスク 1 では「インフルエンザ」の関心が高まってい

\*<sup>8</sup> <http://agora.ex.nii.ac.jp/digital-typhoon/news/2012/TC1204/>





図 15 台風に関連する検索

る時に関心が高まっているアイテムを3つ、実験タスク2では「台風」の関心が高まっている時に関心が高まっているアイテムを3つ発見してもらい、それぞれのアイテムについての理由を回答してもらった。「インフルエンザ」は、毎年1月～2月頃に流行するので、その時期に周期的に関心が高まっている。しかし、2009年4月～5月に新型インフルエンザが世界的に流行しており[37]、この時期にも関心が高まっている。このような非周期的な変動と周期的な変動を組み合わせたクエリを生成することで、「ザナミビル」や「ノイラミニダーゼ阻害薬」のような抗インフルエンザ薬や「グラクソ・スミスクライン」のような製薬会社を発見することができる。

検索タスク(c)は「時期の異なる同一イベントに関連するアイテムの発見」である。オリンピックやワールドカップ等の大きなイベントは様々なアイテムの動向情報に影響を与える。このようなイベントはその開催・発生時期のみに特徴的な変動が発生するためイベントの時期を把握していれば、関連するアイテムを発見することができる。従って、提案する論理演算機能によりクエリを組み合わせることで効率的な探索が期待できる。実験タスク1では2008年、2012年の夏季オリンピックに2大会連続で出場している選手を3人、実験タスク2では2010年、2014年の冬季オリンピックに2大会連続で出場している選手を3人発見してもらった。この時なるべくメダルを獲得している選手を回答してもらうように指示した。

検索タスク(d)は「周期的ではないアイテムの発見」である。検索タスク(a)では周期性のあるアイテムに対する論理演算機能の有効性を検証したが、本タスクでは、非周期性をもつアイテムに

対する論理演算機能の有効性を検証する。非周期性をもつアイテムは、答えが明確ではなく評価が難しいので、正解を用意するのではなく、発見したアイテムをユーザに評価してもらうこととした。実験タスク 1 では 2012 年に周期的ではなく関心が高まっているアイテムを 3 つ、実験タスク 2 では 2013 年に周期的ではなく関心が高まっているアイテムを 3 つ発見してもらい、それぞれのアイテムについての理由を回答してもらった。また、発見したアイテムを実験協力者全員に「アイテムの意外性」、「理由の正当さ・わかりやすさ」の観点からそれぞれ 5 点満点で評価してもらうことで、タスクの達成度を評価した

#### 4.1.3 実験の結果・考察

本項では、論理演算機能を導入したコンテキスト検索エンジンと従来のコンテキスト検索エンジンを用いた実験結果を比較し、論理演算機能の有効性について考察する。

検索タスク (a) に関して、画面遷移数の比較を図 16、検索回数の比較を図 17、実験協力者が検索結果の何ページまで確認したか（確認ページ数）の比較を図 18、解答の正誤と各論理演算子の利用回数を表 7 で示す。本実験では、発見したアイテムのうち 1 つでも正しいものがあった場合を正解とする。

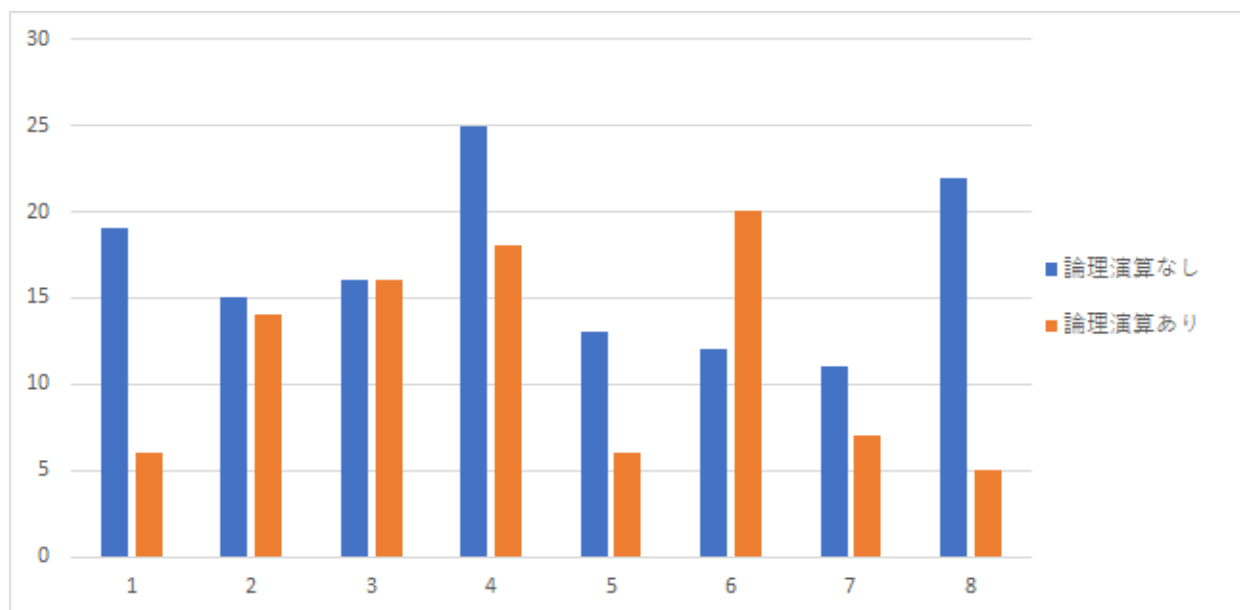


図 16 検索タスク (a) の画面遷移数

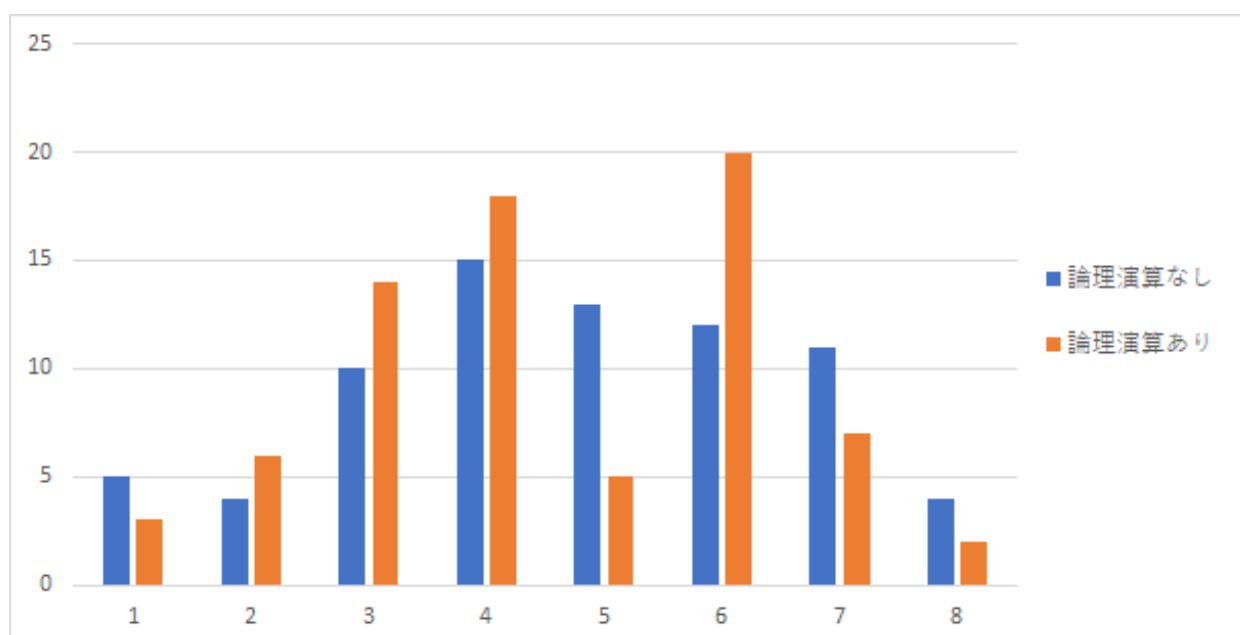


図 17 検索タスク (a) の検索回数

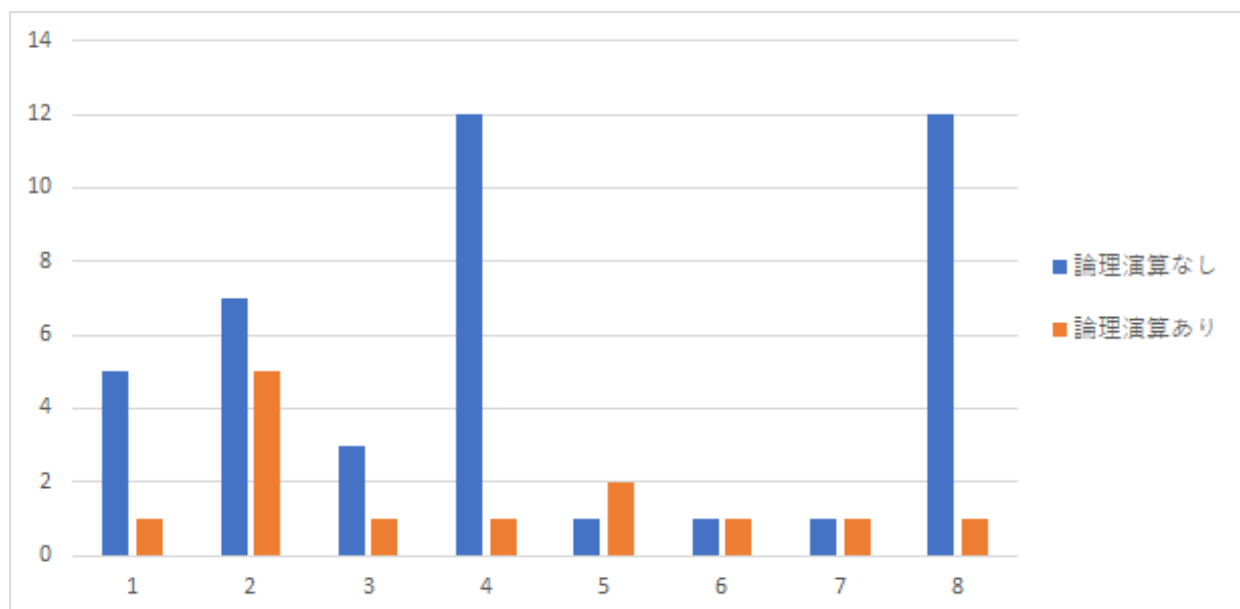


図 18 検索タスク (a) の確認ページ数

表 7 検索タスク (a) の回答の正誤と論理演算子の利用回数

| ID | 論理演算なしの回答 | 論理演算機能ありの回答 | AND | OR | NOT |
|----|-----------|-------------|-----|----|-----|
| 1  | ○         | ○           | 1   | 0  | 0   |
| 2  | ○         | ○           | 2   | 0  | 0   |
| 3  | ○         | ○           | 0   | 0  | 7   |
| 4  | ○         | ○           | 0   | 0  | 0   |
| 5  | ○         | ○           | 3   | 0  | 0   |
| 6  | ×         | ×           | 0   | 5  | 0   |
| 7  | ×         | ×           | 5   | 0  | 0   |
| 8  | ○         | ○           | 1   | 0  | 0   |

表 7 を確認すると、論理演算機能の有無によらず不正解者は 2 名であり、どちらも同じ実験協力者であった。図 16 を確認すると、画面遷移数は ID:3, ID:6 以外の実験協力者は論理演算機能ありの場合の方が少なくなっている。

画面遷移とは検索を行ったり、アイテム名をクリックすることで別の画面に移動することであり、少ない回数で必要な情報を得られる方が効率的な検索が行えていると考える。アイテムの検索時間などは実験に用いた両コンテキスト検索エンジンの中で実装による違いなどがあり、公平な比較ができないため、解答時間ではなく画面遷移数を効率性判断の指標とした。

論理演算機能ありで画面遷移数が減少している 6 名のうち、5 名が論理演算子 AND を利用していた。実験協力者は AND を用いて複数年の 5 月又は 6 月に特徴的変動が発生しているアイテムの検索をすることによって、検索結果を周期的な変動が発生しているアイテムに絞り込んでいた。正解した実験協力者 6 名の画面遷移数について、論理演算機能なしと論理演算機能ありの t 検定を行った結果、有意水準 5 % で統計的に有意な差が確認された。この結果から、論理演算機能を用いることで周期性のあるアイテムを効率的に発見可能と考える。論理演算機能ありで画面遷移数が減少しなかった 2 名は、どちらも AND を利用せず、OR 又は NOT のみを利用していった。その結果、効率的に周期性のあるアイテムの発見ができなかったと考える。ID:4 の実験協力者は論理演算機能を利用していないが、論理演算機能ありの方が画面遷移数が少なくなっている。この実験協力者は、論理演算機能ありの場合に適切な期間で検索を行ったため、論理演算機能を利用しなくても正解を発見できていた。

図 17 を確認すると、検索回数は論理演算機能ありの場合に増えた人 4 名、減った人 4 名であった。この違いは、ユーザのクエリ生成能力に依存している可能性が高いと考える。論理演算機能を導入したことで、複雑なクエリを生成できるようになった。それに伴い、従来のコンテキスト検索エンジンに比べて実験協力者間のクエリ生成能力に差が生じていると考える。ID:6 と ID:8 の実験協力者を例として説明する。ID:6 は OR を用いて 2010 年 5 月に特徴的変動 SI, PEAK, MAX のどれかが発生しているアイテムの検索を行っていた。このような検索では、検索結果が多くなりす

ぎてしまい周期的なアイテムの発見が困難になってしまう。その結果、適切なアイテムが発見できず検索をやり直しており、検索回数が多くなっていた。これに対し、ID:8 は AND を用いて 2010 年 6 月と 2011 年 6 月に特徴的変動 PEAK が発生しているアイテムを検索していた。この検索では、2 年連続で 6 月に PEAK が発生しているアイテムが検索結果として返されており、周期性のあるアイテムに絞り込むことができたため適切なアイテムを発見し検索を終了していた。このように、クエリ生成能力が高い実験協力者は検索回数が少なく、クエリ生成能力が低い実験協力者は検索回数が多くなっていると考えられる。

図 18 を確認すると、確認ページ数は D:5 以外の実験協力者で同じか少なくなっていた。これは論理演算機能によって検索結果が絞り込むことができるので、ランキング下位のアイテムまで確認することなく適切なアイテムを発見可能になったことが要因と考える。

ID:5 の実験協力者は 2011 年から 2014 年の全ての 5 月に特徴的変動 PEAK が発生しているアイテムを検索しており、発見したアイテムは全て検索結果の 1 ページ目にまれていた。正解した実験協力者 6 名の確認ページ数について、論理演算機能なしと論理演算機能ありの t 検定を行った結果、有意水準 10 % で統計的に有意差な差が確認された。

以上のような結果から、論理演算機能を用いて適切なクエリを生成できるユーザにとっては、周期性のあるアイテムの発見に対して論理演算機能は有効と考える。

検索タスク (b) に関して、画面遷移数の比較を図 19、検索回数の比較を図 20、確認ページ数の比較を図 21、解答の正誤と各論理演算子の利用回数を表 8 で示す。

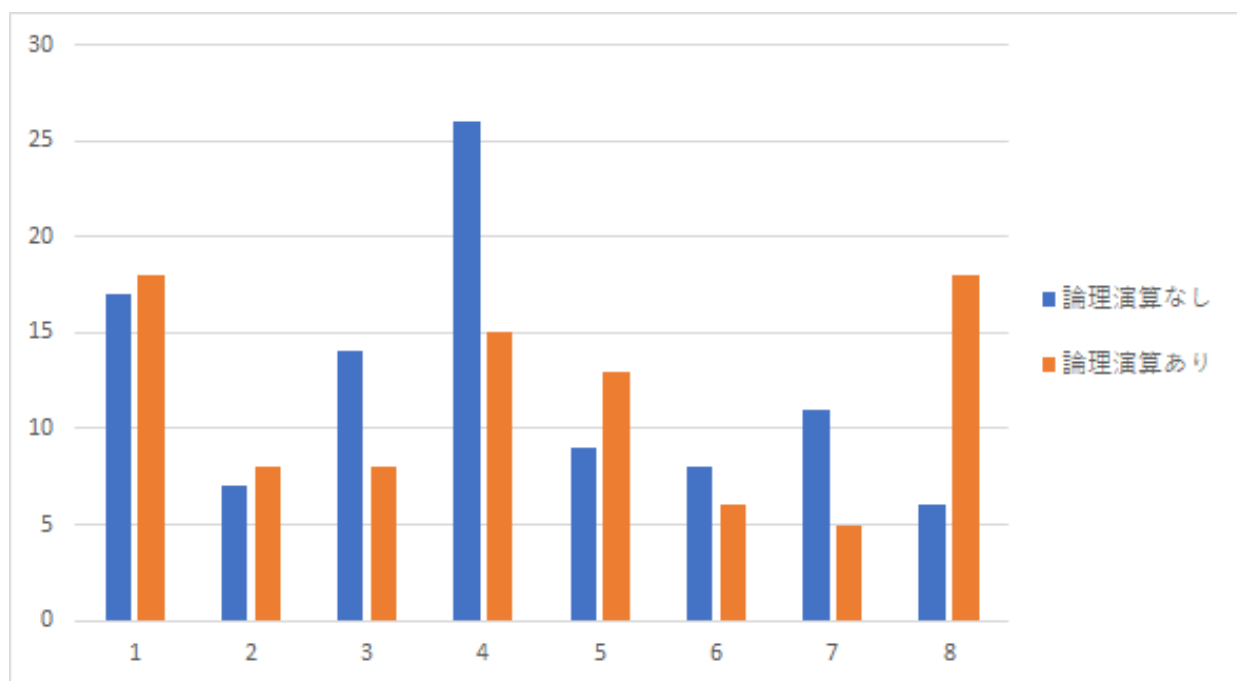


図 19 検索タスク (b) の画面遷移数

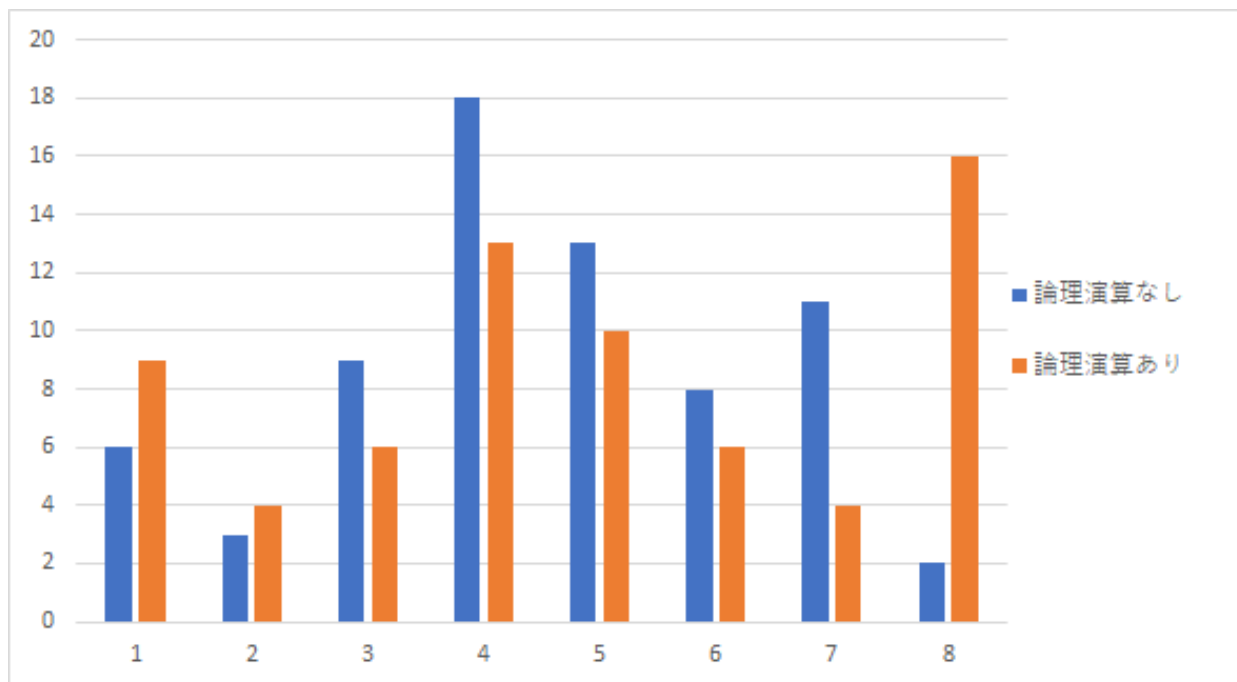


図 20 検索タスク (b) の検索回数

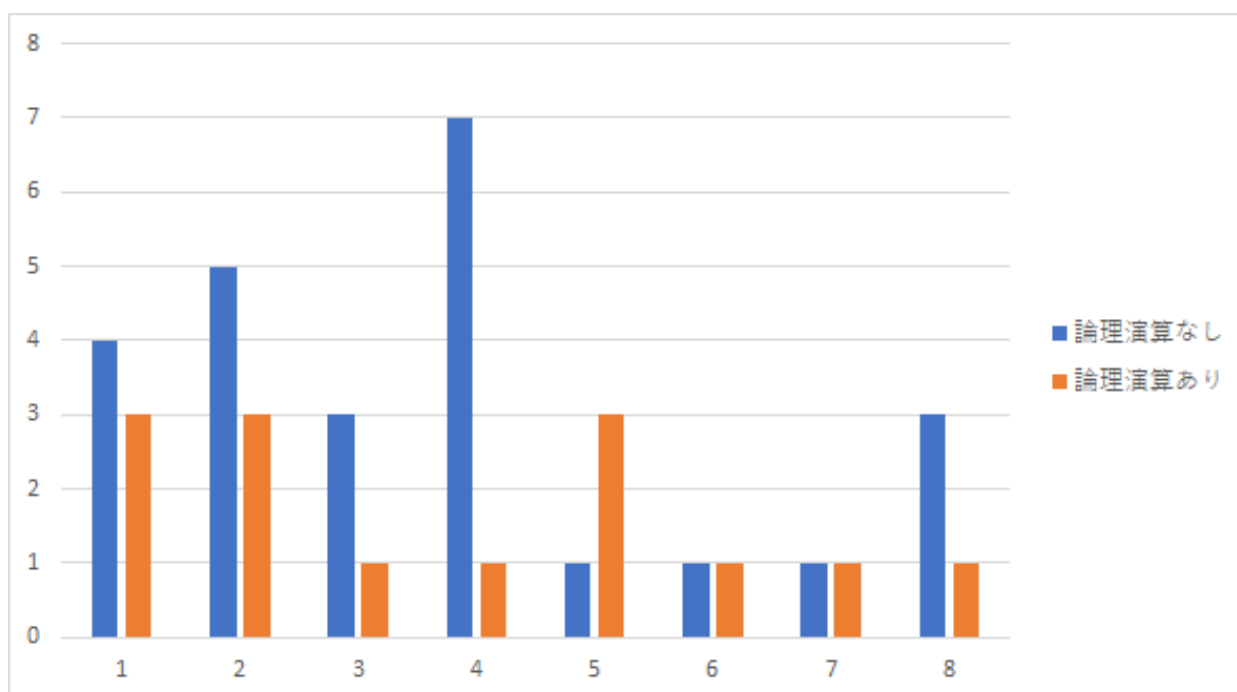


図 21 検索タスク (b) の確認ページ数

表 8 検索タスク (b) の回答の正誤と論理演算子の利用回数

| ID | 論理演算なしの回答 | 論理演算ありの回答 | AND | OR | NOT |
|----|-----------|-----------|-----|----|-----|
| 1  | △         | △         | 2   | 0  | 0   |
| 2  | △         | △         | 1   | 0  | 0   |
| 3  | △         | ×         | 1   | 0  | 0   |
| 4  | ×         | △         | 3   | 0  | 0   |
| 5  | ×         | ×         | 0   | 1  | 0   |
| 6  | ×         | ×         | 1   | 1  | 0   |
| 7  | ×         | ×         | 1   | 0  | 0   |
| 8  | △         | △         | 2   | 0  | 0   |

本タスクにおいて、想定通りの解答をした人はどちらの場合も 0 人であった。そこで、表 8 では各アイテムと周期的な変動が一致しているアイテムを回答した場合には、各アイテムの周期性は理解していると考え△とした。「台風」の場合は「彼岸」や「敬老の日」等の 9 月頃に周期性のあるアイテム、「インフルエンザ」の場合は「正月」や「大学入試センター試験」等の 1 月頃に周期性のあるアイテムを△とした。

論理演算機能なしで△の解答をしている実験協力者は、短い期間を指定することで検索結果の絞り込みを行ったり、特徴的変動 MAX を用いることで検索結果を絞り込みを行った後、検索結果のランキング下位まで確認することでアイテムを発見していた。×となっている実験協力者は、年単位の長い期間を指定しており検索結果が多くなってしまっていた。図 21 をみると、ID:4 の実験協力者以外は検索結果の上位のみを確認したためアイテムを発見できなかったと考える。

論理演算機能ありで△の解答をしている実験参加者の中でも、アイテム発見時に論理演算機能を使っていたのは、ID:1 と ID:8 の 2 名のみであり、短い期間の検索結果を 2 つ用いて絞り込みを行っていた。

ID:3 の実験協力者は論理演算機能ありを先に行っているが、この際に@item opposite を利用しており適切な検索を行えていなかった。一方、論理演算機能なしでは@item opposite を用いず検索を行いアイテムを発見していたため、1 回目の経験を生かし△の解答ができたと考える。

ID:4 の実験協力者は論理演算機能ありで△の解答をしているが、アイテム発見時に論理演算機能を利用していない。また、検索クエリは「台風」、「インフルエンザ」の周期的な変動が起こっている期間に対して特徴的変動 SI が発生しているアイテムを検索している。論理演算機能なしでは、9 月～10 月に周期性のあるアイテムが上位になかったので解答ができなかったと考える。

正しい解答がなかったことから、論理演算機能は検索タスク (b) において適切に利用されなかったことがわかった。原因として、検索タスクの難易度が高かったことが挙げられる。実験協力者は情報検索・動向情報に対して一般的な知識しか有していない。しかし、検索タスク (b) では各アイテムの動向情報を理解し、その動向情報をクエリとして適切に表現、返された検索結果から関係性

を発見しなければならず，コンテキスト検索エンジンでの検索に慣れていない実験協力者には困難であったと考える。

検索タスク (c) に関して，画面遷移数の比較を図 22，検索回数の比較を図 23，確認ページ数の比較を図 24，解答の正誤と各論理演算子の利用回数を表 9 で示す。

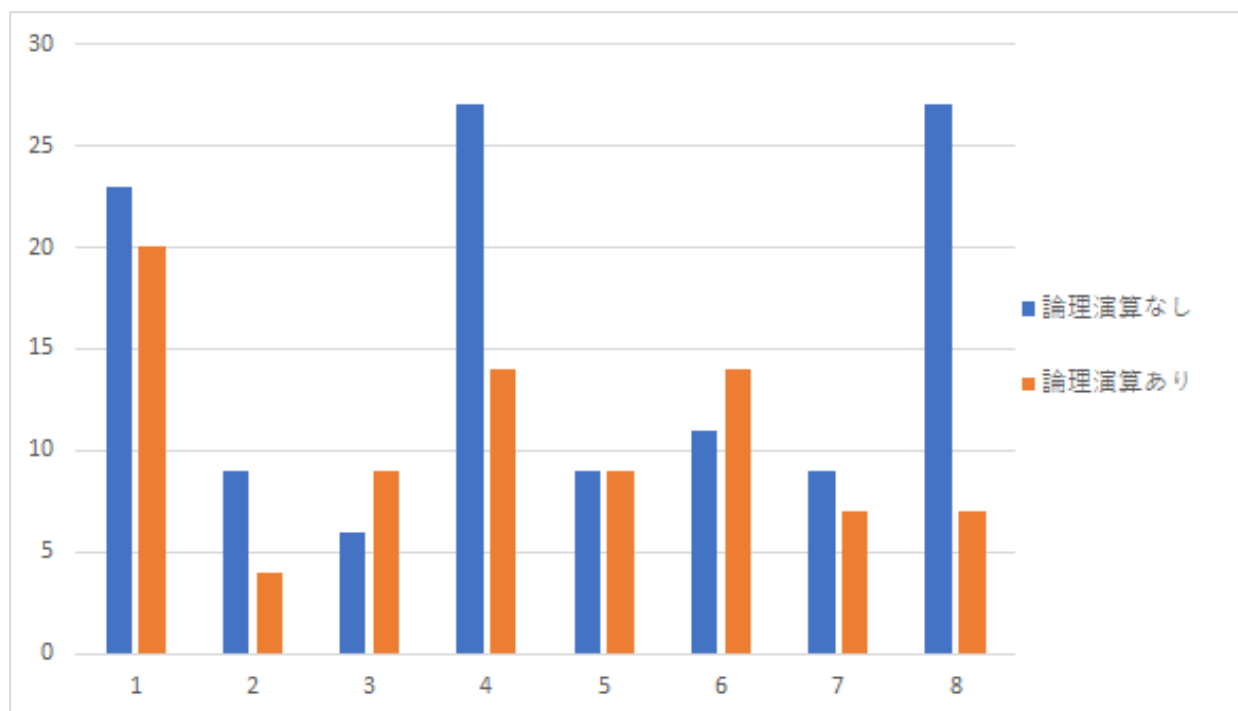


図 22 検索タスク (c) の画面遷移数



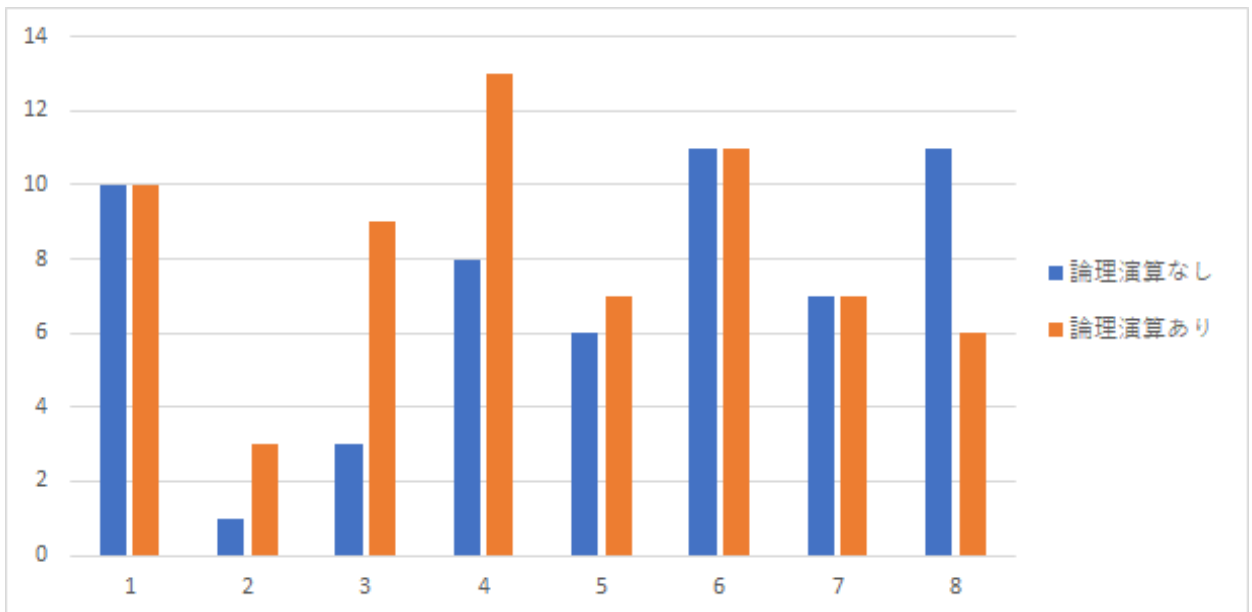


図 23 検索タスク (c) の検索回数

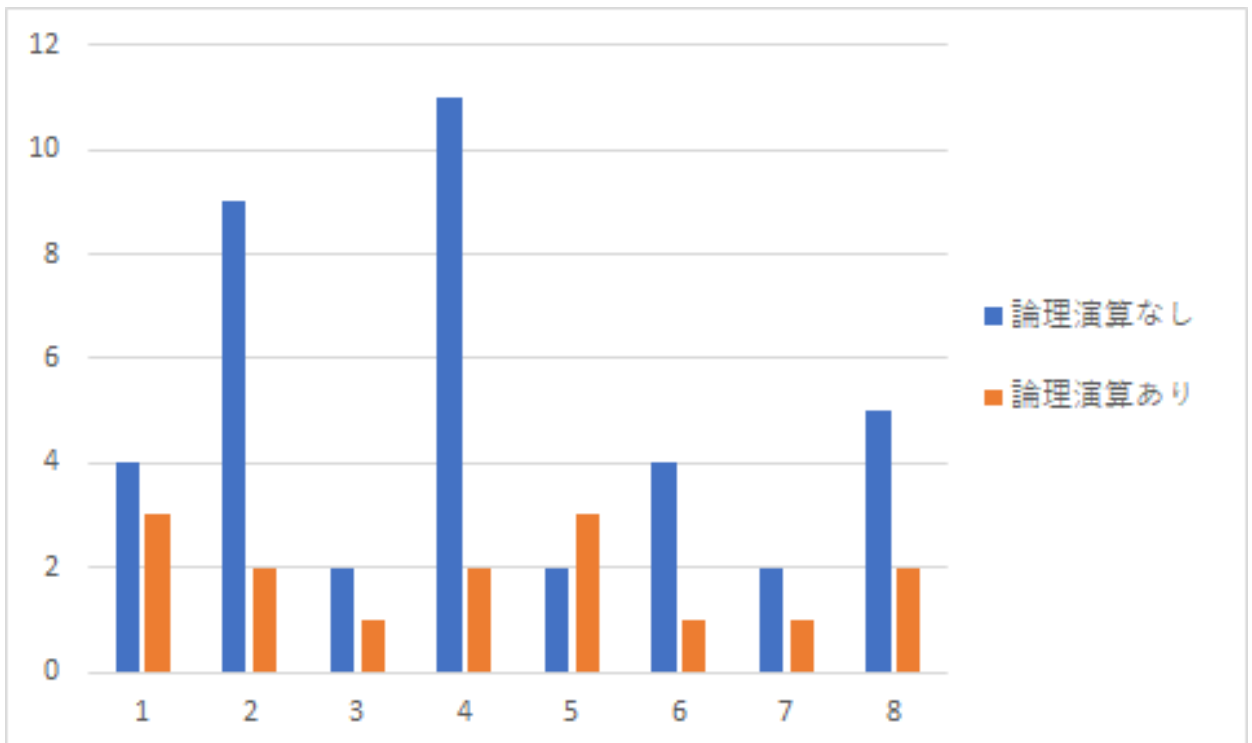


図 24 検索タスク (c) の確認ページ数

表 9 検索タスク (c) の回答の正誤と論理演算子の利用回数

| ID | 論理演算なしの回答 | 論理演算ありの回答 | AND | OR | NOT |
|----|-----------|-----------|-----|----|-----|
| 1  | ×         | ○         | 1   | 0  | 0   |
| 2  | ○         | ○         | 1   | 0  | 0   |
| 3  | ○         | ×         | 3   | 0  | 0   |
| 4  | ○         | ○         | 1   | 0  | 0   |
| 5  | ○         | ○         | 2   | 0  | 0   |
| 6  | ○         | ×         | 2   | 3  | 1   |
| 7  | ○         | ×         | 2   | 0  | 0   |
| 8  | ×         | ×         | 1   | 0  | 0   |

表 9 を確認すると、論理演算機能ありの方が正解している人が少なかった。ID:3 の実験協力者は検索タスク (b) の時と同様に、論理演算機能ありでは@item opposite を利用しており適切な検索を行えておらず、論理演算機能なしでは@item opposite を用いずアイテムを発見していたため、1 回目の経験を生かして正答できたと考える。

ID:6 の実験協力者は検索タスク (a) の時と同様に、OR を用いて 2008 年に特徴的変動 MAX, PEAK が発生しているアイテムを検索していた。このような検索を行うと、期間が長いのでオリンピックに関係のないアイテムが多く返される。論理演算機能なしの場合は複雑なクエリは生成できないので、ランキング下位のアイテムまで確認を行いアイテムを発見していることが図 24 よりわかる。一方、論理演算機能ありの場合は絞り込み方法が適切ではなかったのでアイテムの発見ができなかったと考える。

実験協力者 ID:7, ID:8 は冬季オリンピックの開催時期を夏頃だと勘違いしており、間違った期間で検索を行っていたので適切なアイテムを発見することができなかった。この結果から、正解者数の差は検索タスク間の難易度の違いによるものと考えられる。

図 22 を確認すると、画面遷移数は ID:3, ID:6 の実験協力者のみ、論理演算機能ありの場合の方が多くなっている。この 2 名は論理演算機能ありの時不正解であり、適切なクエリでの絞り込みが行えなかった結果、画面遷移数が増えていると考える。

図 23 を確認すると、検索回数は ID:2, ID:3, ID:4, ID:5 の 4 名が論理演算機能ありの場合の方が増えている。ID:2 の実験協力者は論理演算機能なしでは 1 回しか検索を行っていない。論理演算機能は複数の検索結果を論理演算して行うので、検索を複数回行う必要がある。ID:3 の実験協力者は前述のとおり、クエリの検索結果のタイプで@item opposite を利用して検索を繰り返していた。ID:4 の実験協力者は理由は不明だが全く同じクエリを 7 回検索しており、そのせいで検索回数が多くなっていた。正解できた中で論理演算機能ありの場合に検索回数が増えていないのは ID:1 の実験協力者のみであり、論理演算機能は検索回数の削減に効果はなかったといえる。

図 24 を確認すると、確認ページ数は ID:5 以外の実験協力者で同じか少なくなっていた。これは

論理演算機能によって検索結果が絞り込むことができるので、ランキング下位のアイテムまで確認することなく適切なアイテムを発見することが可能になったためと考える。ID:5の実験協力者は論理演算機能ありの時に3ページ目を確認しているが、2ページ目を確認していないため、操作ミスで2ページ目を確認しようとして3ページ目を確認したと考える。

以上の結果から、時期の異なる同一イベントに関連するアイテムの発見に対して、論理演算機能によって検索回数には大きな変化は見られなかったが、画面遷移数、最高ページ数は減少傾向であるので、効率的な検索結果の絞り込みに対して有効性があると考ええる。

検索タスク (d) についての考察を行う。

論理演算機能なしで発見した2012年のアイテムを表10、論理演算機能ありで発見した2012年のアイテムを表11、論理演算機能なしで発見した2013年のアイテムを表12、論理演算機能ありで発見した2013年のアイテムを表13に示す。全ての表には発見したアイテムに対して「アイテムの意外性」、「理由の正当さ・わかりやすさ」の2つを実験協力者全員が5段階（1が高評価）で評価した結果の平均も示している。

表10 論理演算機能なしで発見された2012年のアイテム

| ID | 発見されたアイテム    | 意外性（平均） | 理由の正当さ・わかりやすさ（平均） |
|----|--------------|---------|-------------------|
| 1  | ソードアートオンライン  | 4       | 1.375             |
| 1  | 平清盛          | 3.875   | 1.5               |
| 1  | 指原莉乃         | 2.125   | 2.125             |
| 2  | 長谷川豊         | 2.375   | 1.875             |
| 2  | GTO_(テレビドラマ) | 2.625   | 1.75              |
| 2  | 日食           | 2.75    | 1.5               |
| 7  | 平清盛          | 3.875   | 2                 |
| 7  | 竹島           | 2.125   | 2                 |
| 8  | ソードアートオンライン  | 4       | 2.375             |
| 8  | ゴールデンボンバー    | 3       | 2.375             |
| 8  | 中二病でも恋がしたい！  | 2.625   | 2.625             |

表 11 論理演算機能ありで発見された 2012 年のアイテム

| ID | 発見されたアイテム        | 意外性（平均） | 理由の正当さ・わかりやすさ（平均） |
|----|------------------|---------|-------------------|
| 3  | 2012 年人類滅亡説      | 3.25    | 1.5               |
| 3  | スギちゃん            | 2.875   | 1.625             |
| 4  | ももいろクローバー Z      | 2.875   | 1.5               |
| 5  | 日食               | 2.75    | 2.375             |
| 5  | 踊る大捜査線 THE MOVIE | 2.875   | 2.375             |
| 6  | 平清盛_(テレビドラマ)     | 3.625   | 1.75              |
| 6  | Another          | 3.25    | 2.125             |
| 6  | ストロベリーナイト        | 3.625   | 1.875             |

表 12 論理演算機能なしで発見された 2013 年のアイテム

| ID | 発見されたアイテム | 意外性（平均） | 理由の正当さ・わかりやすさ（平均） |
|----|-----------|---------|-------------------|
| 3  | 進撃の巨人     | 3.625   | 2                 |
| 3  | 桜塚やっくん    | 3.375   | 1.625             |
| 3  | あまちゃん     | 3.875   | 2                 |
| 4  | 進撃の巨人     | 3.625   | 1.875             |
| 4  | 能年玲奈      | 3.625   | 1.75              |
| 4  | 有村架純      | 2.75    | 3.125             |
| 5  | 桜塚やっくん    | 3.375   | 1.875             |
| 5  | 悪の教典      | 2.75    | 2.125             |
| 5  | 香川照之      | 2.875   | 2.875             |
| 6  | 桜塚やっくん    | 3.375   | 1.375             |
| 6  | あまちゃん     | 3.875   | 1.5               |
| 6  | 能年玲奈      | 3.625   | 1.375             |

表 13 論理演算機能ありで発見された 2013 年のアイテム

| ID | 発見されたアイテム | 意外性（平均） | 理由の正当さ・わかりやすさ（平均） |
|----|-----------|---------|-------------------|
| 8  | 進撃の巨人     | 3.625   | 2.125             |
| 8  | あまちゃん     | 3.875   | 2.125             |
| 8  | 桜塚やっくん    | 3.375   | 2.5               |
| 7  | 進撃の巨人     | 3.625   | 3.875             |
| 7  | あまちゃん     | 3.875   | 3.875             |
| 1  | 桜塚やっくん    | 3.375   | 1.875             |
| 1  | 進撃の巨人     | 3.625   | 2                 |
| 1  | ヨイトマケの歌   | 1.75    | 1.875             |
| 2  | 進撃の巨人     | 3.625   | 1.875             |
| 2  | あまちゃん     | 3.875   | 1.875             |
| 2  | 家族ゲーム     | 2.375   | 1.625             |

2012 年，2013 年で発見したアイテムの多くはその時期に放送されていたドラマ・アニメや芸能人であった．ドラマやアニメは番組が放送されている期間は多く検索され，放送が終わると検索される回数が減少する．ドラマに出演する俳優にも同様の傾向がみられる．また，芸能人の場合はスキャンダルなどが発覚するとその期間のみ多く検索される．また，2.1.3 節で述べた通り，よく検索される物事は，Wikipedia の該当ページもよく閲覧される傾向にある．コンテキスト検索エンジンでは，動向情報の変動の激しさを用いてランキングスコアを生成しているので，このように短期間で多く検索され，放送が終わると検索回数が減少するアニメやドラマ，映画等の WikipediaPageView などは検索結果の上位になりやすく，この検索タスクで多く発見されていると考える．

表 13 の「ヨイトマケの歌」は全実験結果の中で意外性が非常に高く評価されているが，アイテム発見時に論理演算機能を利用していない．また，意外性の評価が比較的に高い表 10 の「指原莉乃」や「竹島」は論理演算機能なしで発見されており，意外性のあるアイテムに対して論理演算機能は効果はなかったといえる．理由の正当さ・わかりやすさの評価はアイテムに対する説明を詳細に書いてあるアイテムの場合に高く，実験協力者に依存していると考ええる．

この検索タスクでの，論理演算子の利用回数を表 14 で示す．

表 14 検索タスク (d) での論理演算子の利用回数

| ID | AND | OR | NOT |
|----|-----|----|-----|
| 1  | 0   | 0  | 0   |
| 2  | 2   | 0  | 0   |
| 3  | 0   | 0  | 2   |
| 4  | 0   | 0  | 0   |
| 5  | 0   | 2  | 1   |
| 6  | 4   | 8  | 1   |
| 7  | 2   | 0  | 0   |
| 8  | 0   | 0  | 0   |

今までの検索タスクと比較すると、NOT 演算子を利用している実験協力者が多いことがわかった。これは、NOT 演算子を用いることで他の年に特徴的変動が発生していないアイテムを検索することで、周期的なアイテムを検索結果から取り除こうとしたためである。

2012 年に発見したアイテムの、意外性の平均は論理演算機能なしでは 3.03、論理演算機能ありでは 3.14、理由の正当さ・わかりやすさの平均は論理演算機能なしでは 1.95、論理演算機能ありでは 1.89 であり、論理演算機能による差はみられなかった。2013 年に発見したアイテムの、意外性の平均は論理演算機能なしは 3.70、論理演算機能ありでは 3.36、理由の正当さ・わかりやすさの平均は論理演算機能なしでは 1.96、論理演算機能ありでは 2.33 であり、こちらも論理演算機能による差はみられなかった。

以上の結果から、非周期的なアイテムの発見において論理演算機能は有効性がみられなかった。このタスクにおいて動向情報の変動の激しさを用いたランキングの有効性が高く、論理演算機能を使う必要があまりなかったことが原因と考える。

実験後アンケートの回答結果を図 25, 26, 27 に示す。図 25 は論理演算機能のわかりやすさを 5 段階で評価してもらった。図 26, 27 ではコンテキスト検索エンジンの速度について論理演算機能なしとあり場合それぞれについて、5 段階（1 が高評価）で評価してもらった。

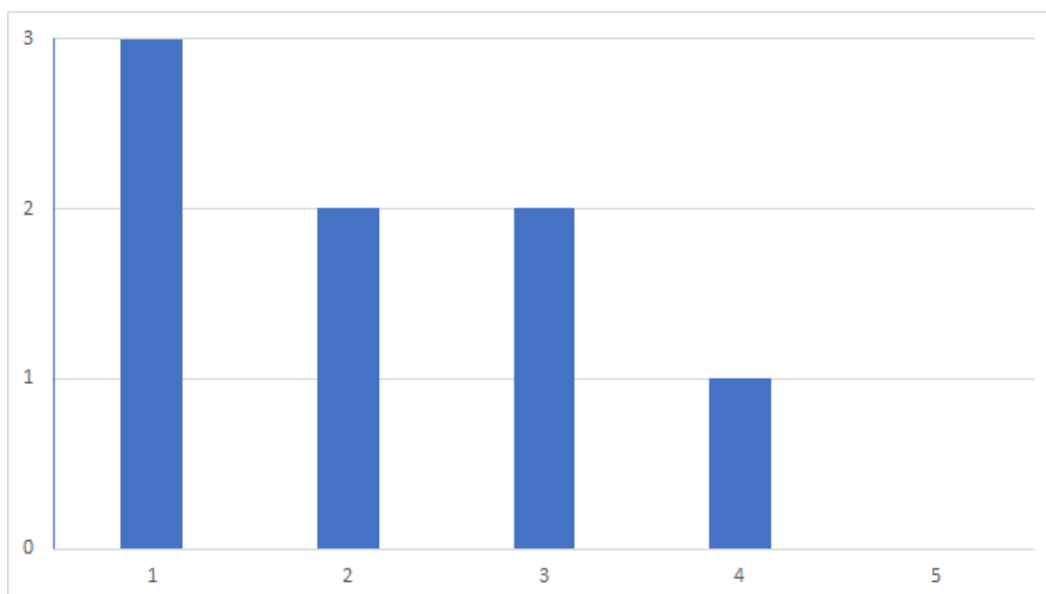


図 25 アンケート回答:論理演算機能はわかりやすかったか? (1:わかりやすい~5:わかりにくい)

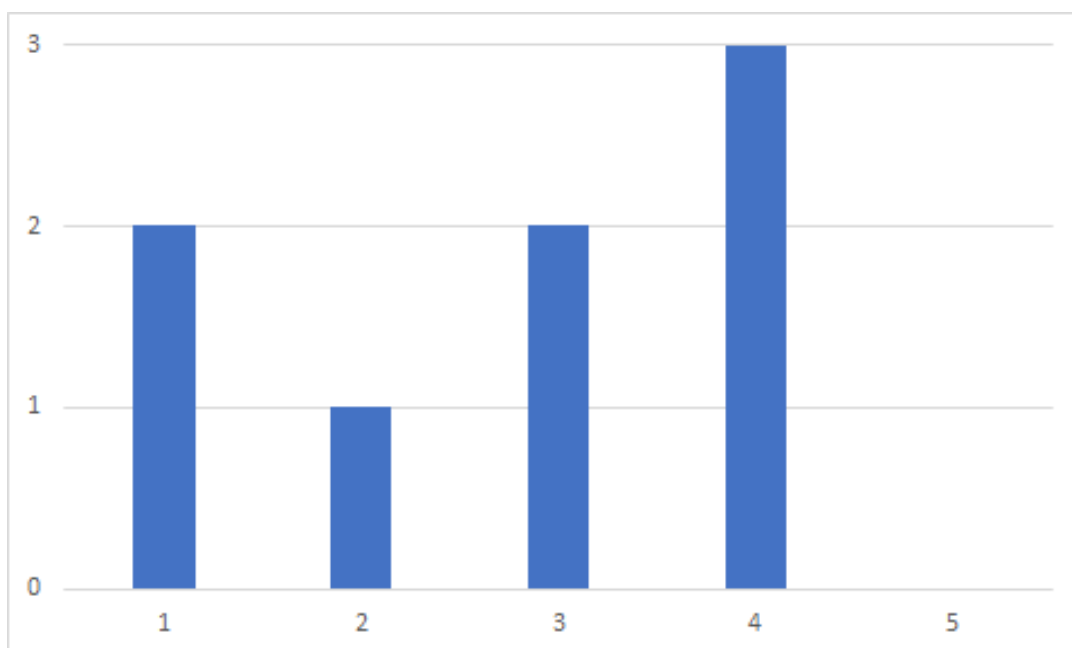


図 26 アンケート回答:論理演算機能なしのコンテキスト検索エンジンの速度 (1:気にならない~5:遅い)

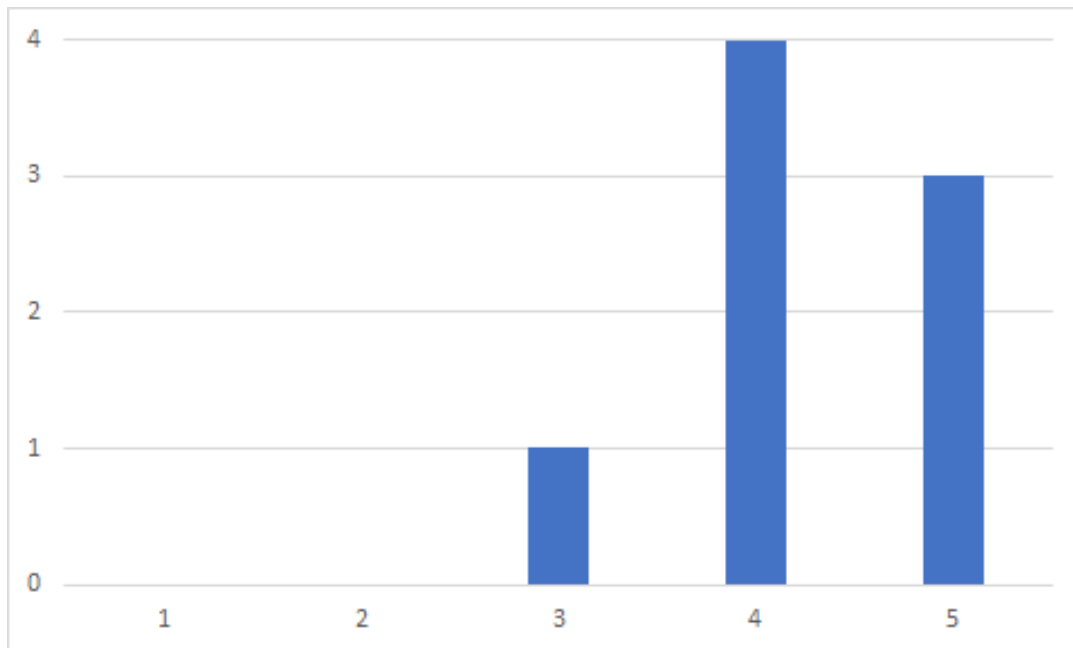


図 27 アンケート回答:論理演算機能ありのコンテキスト検索エンジンの速度 (1:気にならない~5:遅い)

図 25 の結果から、論理演算機能自体は実験協力者に理解されているといえる。図 26、図 27 を確認すると、論理演算機能なしの場合では実験協力者によって評価が分かれたが、論理演算機能ありの場合は遅いと回答する実験協力者が多かった。論理演算機能自体は検索結果の絞り込みに有効ではあるが、速度においてユーザの負担が大きいのことがわかる。

実験中に編集ページを利用した実験参加者は 3 名のみであり、あまり利用されていないことがわかった。クエリ編集ページに移動する必要があったことが原因の一つと考える。

論理演算機能をどのように利用したかに対するアンケートについては、複数の期間に対して AND 検索を行ったとの回答が 5 名からあったことから、周期性のあるアイテムの発見を目的としたタスク (a) や特徴的変動の発生時期がわかっているアイテムの発見を目的とした検索タスク (c) において有効性があったと考える。それ以外の利用方法として、検索タスク (d) において、他の期間の検索結果に対して NOT を用いることで周期的なアイテムを検索結果から取り除こうとした実験協力者が 2 名、OR を用いて同一期間で異なる特徴的変動の検索結果を足し合わせることで、アイテムに対する特徴的変動を漏らさないようにした実験協力者が 1 名いた。実験前に読んでもらったガイドラインについては好評であり、実験協力者のシステム理解に有効と考える。



## 4.2 予備実験

本節では、ユーザの様々な検索行動を収集するため、ユーザの興味によって検索タスクを決定する、従来の Web 検索エンジンの利用に近いタスクを想定した予備実験の結果を示す。

### 4.2.1 実験概要

工学系大学生・大学院生 11 名に対し、政治関係のネット記事<sup>\*9</sup>を読んでもらいその記事から関連するアイテムおよびその理由を回答してもらった。この記事を選んだ理由として、民主党政権の約 3 年 4 ヶ月間という長い期間についてまとめられている点、この期間内に政権交代が起こった 2 回の衆議院選挙、1 回の参議院選挙や、2011 年 3 月 11 日に発生した東日本大震災という他のアイテムに大きな影響を与えるイベントがある点、政治関連の知識は一般に知られており、事前知識による実験協力者の能力の差が少ない点があげられる。この実験に用いたネット記事を付録 C で示す。各実験協力者には、3.2 節で作成したコンテキスト検索エンジンのガイドラインを読んでもらった後、記事内容に関係するアイテムの発見を行ってもらった。実験の制限時間は 30 分に設定し、制限時間内になるべく多くのアイテムを発見してもらうように指示した。また、多様な観点から検索を行う動機づけとして簡単な景品を用意し、他の人の発見したアイテムとかぶらないアイテムを発見した場合に贈呈することとした。従来の Web 検索エンジンの利用は、コンテキスト検索エンジンで発見したアイテムに関する情報を検索する場合のみ許可した。実験後には、検索行動に関するアンケートに回答してもらった。4.1 節の実験でコンテキスト検索エンジンの論理演算機能の速度が遅いとの評価であった。そこで、本実験前に速度に対する改良を行ったため、速度に関する質問も実験後のアンケートに組み込んだ。

### 4.2.2 実験の結果・考察

本実験で発見されたアイテムと、回答した実験協力者を表 15 で示す。表 15 で赤字のアイテムは実験協力者が発見したアイテムの中で、事前に記事との関連に気付いていなかった、または意外なアイテムと回答したアイテムである。青字のアイテムは記事との関連が薄いアイテムである。

---

<sup>\*9</sup> <https://www.jiji.com/jc/graphics?p=ve.pol.election-syugin20121124j-04-w680>

表 15 発見されたアイテムとそのアイテムを発見した実験協力者

| 実験協力者 ID | 発見したアイテム        |
|----------|-----------------|
| 1,5      | スリーマイル島原子力発電所事故 |
| 1,5      | チェルノブイリ原子力発電所事故 |
| 2        | 福島第一原子力発電所事故    |
| 3.11     | 内閣総理大臣の一覧       |
| 4        | 公明党             |
| 4        | 国民新党            |
| 4        | 村山内閣            |
| 4        | 田中角栄            |
| 4        | 宏池会_(谷垣派)       |
| 5        | 高井真一            |
| 5        | 鳩山邦夫            |
| 5        | 杉村太蔵            |
| 5,7      | 炉心融解            |
| 7        | 放射能             |
| 8        | M1 グランプリ        |
| 8        | ファイナルファンタジー XII |
| 8        | プレイステーション 3     |
| 8        | 2009 年          |
| 9        | 麻生太郎            |
| 11       | 野口美佳            |
| 11       | 官僚たちの夏          |

| 実験協力者 ID | 発見したアイテム |
|----------|----------|
| 1        | 坂の上の雲    |
| 2        | 普天間飛行場   |
| 2,5      | 福島瑞穂     |
| 3        | 日韓通貨スワップ |
| 4        | 田中直紀     |
| 4        | 政権交代     |
| 4        | 安定多数     |
| 4        | 東京 18 区  |
| 5,7      | 小沢一郎     |
| 5        | 鳩山一郎     |
| 5        | 鳩山紀一郎    |
| 5        | 赤松広隆     |
| 6        | 国内総生産    |
| 7,9      | 選挙速報     |
| 8        | 秋山真之     |
| 8        | 日清戦争     |
| 8        | 鳥人       |
| 8        | 共産主義     |
| 9        | 原発       |
| 11       | 大原麗子     |

実験協力者の多くは記事内から鳩山由紀夫、菅直人等の人物名、尖閣諸島や普天間飛行場等の地名、ねじれ国会や消費増税等の用語をクエリとして検索を行っており、政治家や政治に関係のあるアイテムが多く発見されている。また、記事内の期間に発生した東日本大震災に関係するアイテムとして、「放射能」や「炉心融解」、類似事件として「チェルノブイリ原子力発電所事故」や「スリーマイル島原子力発電所事故」を発見している。一方で、「M1 グランプリ」、「大原麗子」などの同一期間に同じ特徴的変動が発生しているだけの関係の薄いアイテムも多く発見していた。

発見した 48 アイテムのうち 23 アイテムにおいて、アイテム発見時に論理演算機能を利用していなかった。13 アイテムが 2 回の検索結果に対して論理演算機能を用いて、12 アイテムが 3 回の検索結果に対して論理演算機能を用いて発見していた。また、ID:8 の実験協力者は 3 回の検索結果に対して論理演算機能を用いた検索結果から全てのアイテムを発見していたが、1 回目の検索結果

から検索結果件数が減っていなかったため、論理演算機能を有効に利用できていなかった。  
実験協力者が検索結果の何ページ目でアイテムを発見したかを図 28 に示す。

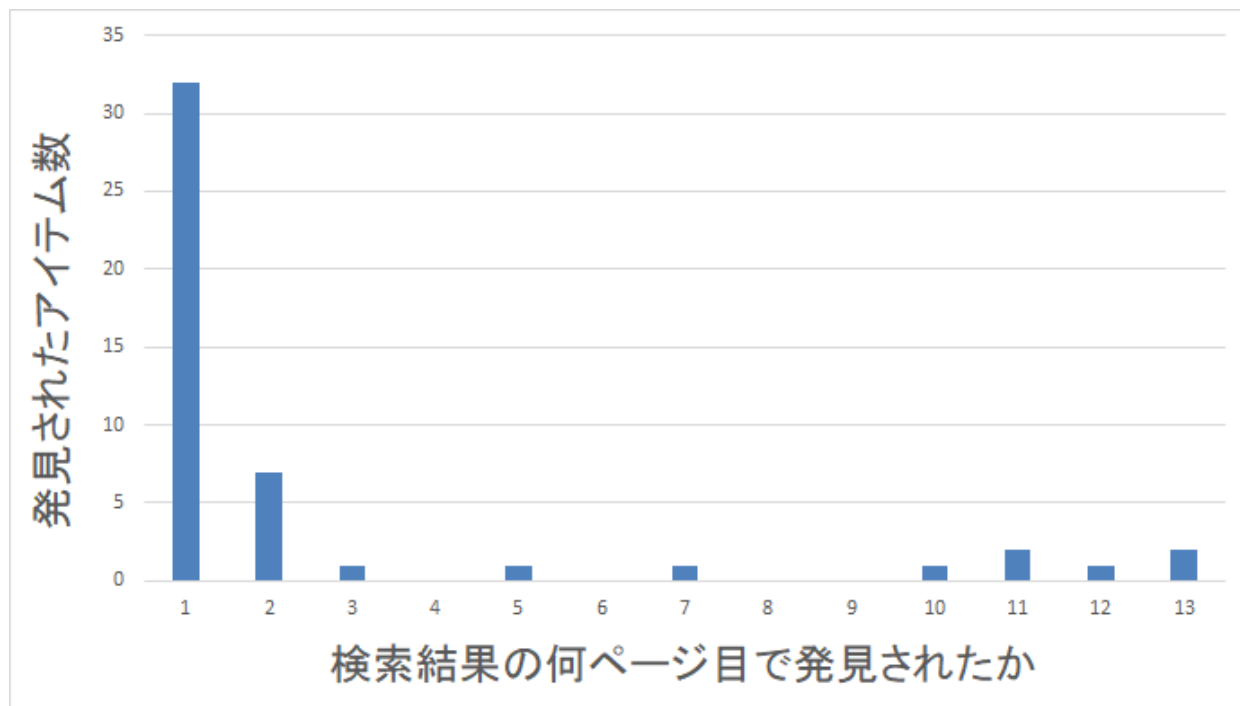


図 28 実験協力者がアイテムを発見したページ

図 28 を確認すると、最初の 1, 2 ページで発見されたアイテムが多いことがわかる。また、7 ページ目以降でアイテムを発見している実験協力者は ID:4 のみであった。2 つ程度のクエリで検索を行い、結果は最初の 1, 2 ページのみを確認するという検索行動は 2.3 節で述べた、Web 検索エンジンでの典型的な情報探索行動と一致している。

実験協力者の特徴的変動の利用回数を表 16、論理演算子の利用回数を表 17 で示す。

表 16 特徴的変動の利用回数

| 実験協力者 ID | ALL | MAX | MIN | SI | SD | PEAK | BOTTOM | NONE |
|----------|-----|-----|-----|----|----|------|--------|------|
| 1        | 2   | 45  | 0   | 5  | 0  | 5    | 0      | 0    |
| 2        | 1   | 42  | 1   | 23 | 0  | 20   | 1      | 0    |
| 3        | 4   | 26  | 0   | 16 | 0  | 5    | 0      | 0    |
| 4        | 0   | 2   | 0   | 6  | 0  | 3    | 0      | 0    |
| 5        | 5   | 9   | 0   | 0  | 0  | 13   | 0      | 0    |
| 6        | 1   | 21  | 0   | 18 | 0  | 1    | 0      | 0    |
| 7        | 0   | 12  | 0   | 8  | 0  | 2    | 0      | 0    |
| 8        | 3   | 12  | 0   | 0  | 0  | 0    | 0      | 0    |
| 9        | 0   | 17  | 0   | 5  | 0  | 19   | 0      | 0    |
| 10       | 4   | 2   | 1   | 0  | 0  | 26   | 2      | 0    |
| 11       | 0   | 0   | 0   | 5  | 0  | 10   | 0      | 0    |
| 計        | 20  | 188 | 2   | 86 | 0  | 104  | 3      | 0    |

表 17 論理演算子の利用回数

| 実験協力者 ID | 利用なし | AND | OR | NOT |
|----------|------|-----|----|-----|
| 1        | 24   | 22  | 6  | 5   |
| 2        | 62   | 19  | 7  | 0   |
| 3        | 39   | 11  | 1  | 0   |
| 4        | 7    | 4   | 0  | 0   |
| 5        | 19   | 8   | 0  | 0   |
| 6        | 22   | 18  | 0  | 0   |
| 7        | 20   | 0   | 2  | 0   |
| 8        | 10   | 5   | 0  | 0   |
| 9        | 39   | 2   | 0  | 0   |
| 10       | 23   | 12  | 0  | 0   |
| 11       | 10   | 5   | 0  | 0   |
| 計        | 275  | 106 | 17 | 5   |

特徴的変動は MAX, PEAK, SI の順に利用回数が多かった。これは、最大値や山、急上昇などの方が時系列データの特徴として直感的にわかりやすく利用しやすかったためと考える。

論理演算機能において、論理演算機能を利用した場合の 8 割以上が AND であった。AND 演算子は、従来の Web 検索エンジンでの絞り込みで最も使われており、実験協力者にとって理解しやすく効率的な絞り込みが行えるため利用されたと考える。

4.1 節の実験では編集ページを用いたクエリの編集を行った実験協力者は少なかったが、本実験のログを確認すると、8 名がクエリの編集を行っており、そのうち 7 名が検索結果画面で編集を行っていた。従って、検索結果画面にクエリ編集機能を導入した効果が認められる。

実験協力者のログを確認すると、4 つの点でコンテキスト検索エンジンのクエリ生成に対して誤った認識をしている可能性を発見した。

1 点目は検索結果のタイプ選択である。コンテキスト検索エンジンでは、指定する検索タイプに応じて、クエリとして指定する必要があるものが異なる。検索結果のタイプが@period の場合、アイテム名の指定が必須であり、検索結果は指定したアイテムの期間のみになる。@item の場合、期間かアイテム名の指定が必須であり、指定した期間内で特徴的変動が発生したアイテム、あるいは指定アイテムで特徴的変動が発生した期間のどこか 1 か所以上で特徴的変動が発生したアイテムを検索結果として返す。しかし、実験協力者の中には、@period と@item を指定したクエリを AND で組み合わせたり、@period 同士を AND で組み合わせたりしていたが、この場合意味のある絞り込みは行えていないことになる。

2 点目はアイテム名のみを指定した検索である。実験協力者のクエリログを確認すると、アイテム名のみを指定した検索が多くみられた。この場合、当該アイテムの特徴的変動が発生している全期間が検索対象となる。MAX や MIN は各アイテムにつき該当期間は 1 つしか存在しないが、SI, PEAK などは該当期間が複数存在するため、アイテムによっては検索結果が多くなってしまい適切な絞り込みが行えない。

3 点目は特徴的変動 MAX, MIN と論理演算機能の組み合わせ方である。実験協力者の中には、複数の期間又はアイテムの MAX を指定した検索結果を AND で組み合わせている。しかし、上記で説明した通り MAX, MIN は各アイテムで該当期間が一つしかないため、組み合わせると検索結果が 0 件となる場合が多い。さらに、実験協力者は、後述するアイテム名での部分一致検索と併用することで複数のアイテム名の MAX が発生した期間で検索を行うことがあり、その場合は適切な検索を行えていない。

4 点目はアイテム名の部分一致検索である。部分一致検索はアイテム名の表記ゆれに対応しており、検索結果のタイプが@period の場合、適切なアイテム名の選定に有効な手段である。例えば、ジャニーズの嵐を検索したい場合、アイテム名に嵐を指定して完全一致検索を行うと自然現象の嵐が検索されてしまう。ここでアイテム名に嵐を指定して部分一致検索を行い、検索結果を確認すると嵐\_\_（ジャニーズ）というアイテムを発見することができ、適切なアイテム名を知ることができる。しかし、検索結果のタイプが@item の場合、複数のアイテム名で特徴的変動が発生した期間を検索対象とするため検索結果が多くなってしまう。

これらの 4 点に対して正しい知識を、3.3 節で述べた「コンテキスト検索エンジンの検索システ

ムについての知識」とした。

実験後のアンケートからコンテキスト検索エンジンの速度に関する回答を図 29 に示す。図 29 の結果から、改良の効果が認められ、ユーザの負担が小さくなったことがわかる。

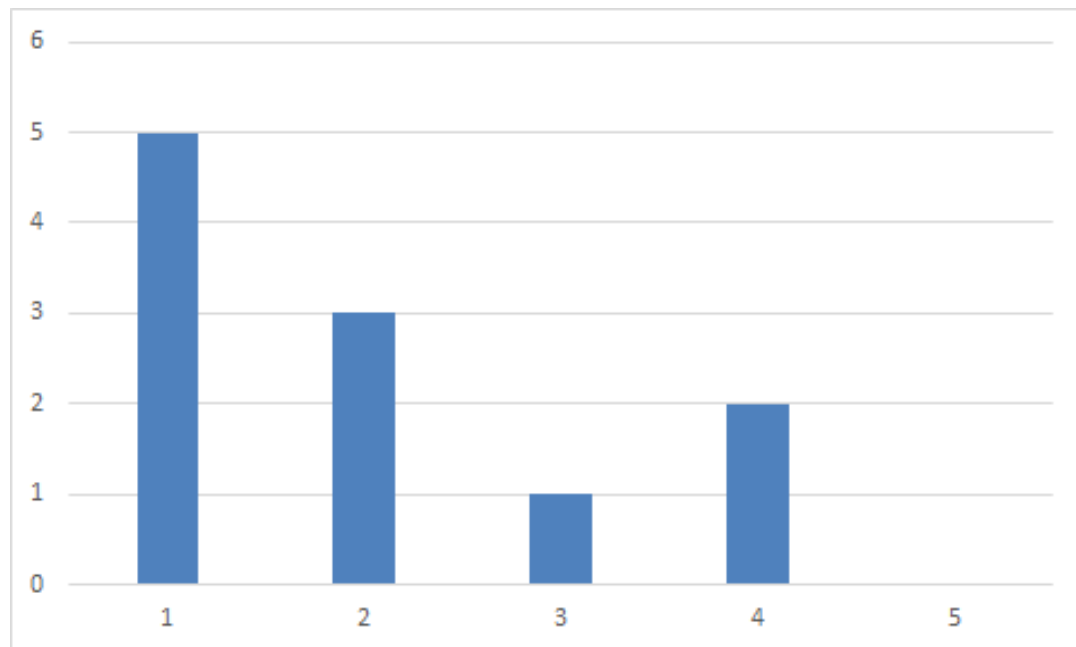


図 29 アンケート回答:コンテキスト検索エンジンの速度 (1:気にならない~5:遅い)

本実験で事前に記事との関連に気付いていなかった、または意外なアイテムを発見できたと回答した実験協力者は7名であった。どのような検索を行ったかに対するアンケートについては、実験参加者のほとんどが記事内のアイテム名を用いた検索のみを行っており、期間を指定したと回答した実験参加者は3名のみであった。また、政治関連のアイテムはあまり興味がなくアイテムを探すのが難しかったとの感想があった。

## 4.3 効率的な検索行動に対する評価実験

本節では、3.3 節で説明したコンテキスト検索エンジンの効率的な検索行動に対する知識をユーザに提示したうえで行った評価実験の結果を示し、ユーザに対する効率的な検索行動支援の有効性について考察する。

### 4.3.1 実験概要

本実験では、4.2 節の予備実験と同じ実験協力者 11 名に対し、2 つのタスクを行ってもらった。各実験協力者には、3.3 節で作成したコンテキスト検索エンジンにおける効率的な検索行動に対する知識を読んでもらった後、タスクを行ってもらう。タスク 1 では、予備実験と同じタスクを行ってもらった。この際に、制限時間を予備実験より 10 分短い 20 分に設定した。タスク 2 では、予備実験の感想から政治以外の検索タスクとして自分の好きなものに関連するアイテムおよびその理由を回答してもらった。ただし好きなものはコンテキスト検索エンジン内にアイテムとして存在するものに限定した。タスク 2 は、制限時間は 10 分に設定した。従来の Web 検索エンジンの利用は、コンテキスト検索エンジンで発見したアイテムに関する情報を検索する場合のみ許可した。実験後には、効率的な検索行動に関するアンケートに回答してもらった。

### 4.3.2 実験の結果・考察

本実験のタスク 1 で発見されたアイテムと、回答した実験協力者を表 18 で示す。表 18 で赤字のアイテムは実験協力者が発見したアイテムの中で、事前に記事との関連に気づいていなかった、または意外なアイテムと回答したアイテムである。青字のアイテムは記事との関連が薄いアイテムである。

表 18 タスク 1 で発見されたアイテムとそのアイテムを発見した実験協力者

| 実験協力者 ID | 発見したアイテム       |
|----------|----------------|
| 1,4,5    | 田中美絵子          |
| 2,10     | 鳩山邦夫           |
| 3        | 虚言癖            |
| 4        | 自由民主党          |
| 4        | 日本の政党一覧        |
| 4        | 右翼             |
| 4        | 赤松広隆           |
| 4        | 原口一博           |
| 4        | 蓮舫             |
| 5        | 福島瑞穂           |
| 6        | 藤井裕久           |
| 6        | 国会議事堂          |
| 6        | 日本の行政機関        |
| 6        | 自由民主党幹事長       |
| 6        | 青木幹雄           |
| 7        | 選挙開票速報         |
| 7        | 中川昭一           |
| 11       | 第 46 回衆議院議員総選挙 |
| 11       | 戦争             |

| 実験協力者 ID | 発見したアイテム         |
|----------|------------------|
| 1        | 盛土               |
| 3        | 福田組              |
| 4        | 民主党代表            |
| 4        | 小泉劇場             |
| 4        | 谷垣禎一             |
| 4,8      | 社会民主党_(日本_1996-) |
| 4,9      | 民主党国会議員一覧        |
| 4        | 菅直人内閣の政策         |
| 4        | 東北地方太平洋沖地震       |
| 5        | 尖閣諸島領有権問題        |
| 6        | 平野博文             |
| 6        | 細川護熙             |
| 6        | 小沢重喜             |
| 6        | 自由民主党国会議員一覧      |
| 6        | 国家公安委員会委員長       |
| 7        | 鳩山由紀夫            |
| 9        | 選挙速報             |
| 11       | 任期               |

実験協力者の多くは前回と同様に記事内のアイテムをクエリとして検索を行っているが、前回の検索と異なりアイテム名のみではなく同時に期間も指定した検索が多くなっており、効率的な検索結果の絞り込みが行われていた。アイテム名が指定されている@item 検索のうち、期間が指定されている検索の比率は、予備実験では約 38 %であったのに対し、本実験では約 62 %であった。この結果は、4.2 節で述べたアイテム名のみを指定した検索で提供した知識が理解されたことが要因と考える。また、予備実験では見られた同一期間に同じ特徴的変動が発生しているだけで関係の薄いアイテムが今回の実験ではほとんどなくなっていた。

発見した 42 アイテムのうちアイテム発見時に論理演算機能を利用していないものは 14 アイテムであった。24 アイテムは 2 回の検索結果に対して論理演算機能を用いて、2 アイテムが 3 回の検索結果に対して論理演算機能を用いて、2 アイテムが 4 回の検索結果に対して論理演算機能を用いて発見していた。予備実験では、論理演算機能を用いない検索で発見されるアイテムが 1 番多かったのに対し、本実験では、2 つの検索結果を絞り込んだ検索結果から発見されるアイテムが 1



多かった。これは、効率的な検索行動に対する知識を提示したことによって、複雑な検索クエリを利用して検索を行ったと考える。

実験協力者がタスク 1 において何ページ目でアイテムを発見したかを図 30 で示す。

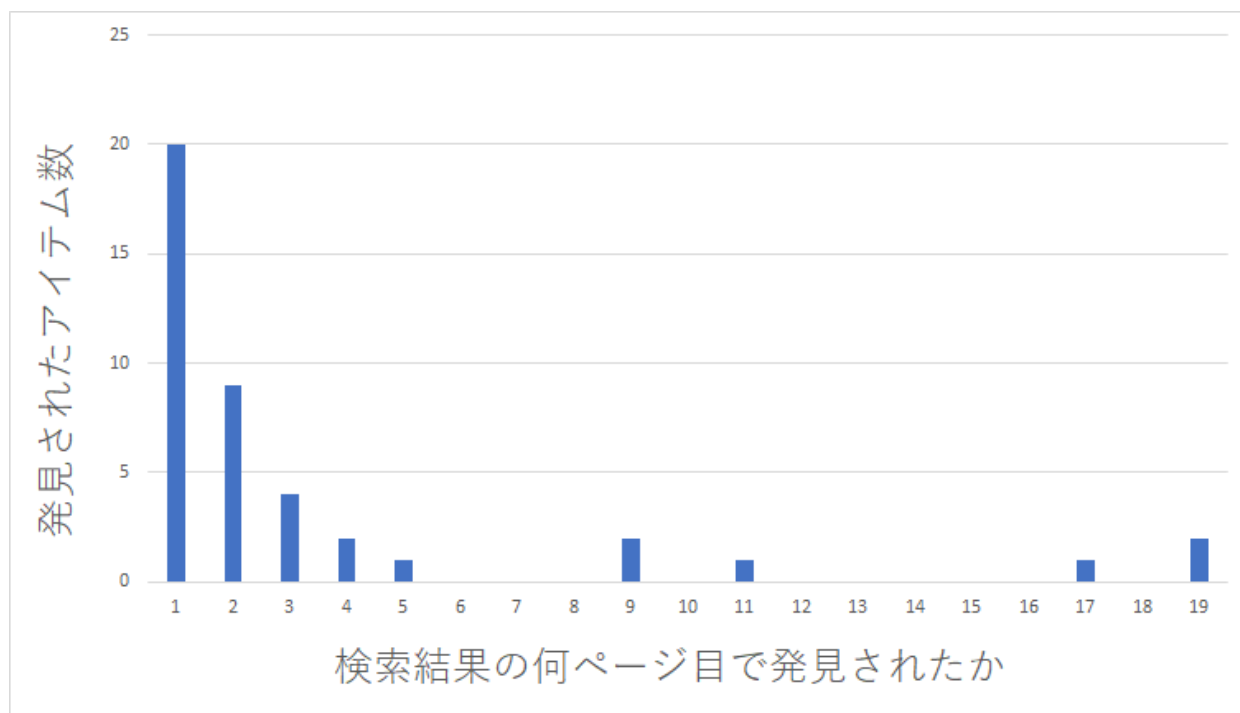


図 30 実験協力者がアイテムを発見したページ

図 30 を確認すると、最初の 1 ページ目で発見されたアイテムが 1 番多かった。しかし、半分以上のアイテムが 2 ページ目以降で発見されており、予備実験と比較して実験協力者がランキング下位のアイテムまで確認したことがわかる。また、5 ページ目以降でアイテムを発見している実験協力者は ID:4 と ID:7 の 2 名のみであった。この結果は、3.3 節で述べた網羅的な検索結果の確認で提供した知識によるものと考ええる。

タスク 2 で実験協力者が関連を調べたアイテムと、発見されたアイテムを表 19 で示す。表 19 で赤字のアイテムは実験協力者が発見したアイテムの中で、事前に記事との関連に気づいていなかった、または意外なアイテムと回答したアイテムである。

表 19 発見されたアイテムとそのアイテムを発見した実験協力者

| 実験協力者 ID | 調べたアイテム          | 発見したアイテム        |
|----------|------------------|-----------------|
| 1        | サッカー             | イナズマイレブン        |
| 1        | サッカー             | フィリップトルシエ       |
| 2        | ギター              | けいおん!           |
| 2        | ギター              | X JAPAN         |
| 2        | ギター              | 竹達彩奈            |
| 2        | ギター              | モーニング娘。         |
| 2        | ギター              | 福山雅治            |
| 2        | ギター              | Hide            |
| 3        | 旅行・海外旅行          | 日本航空 123 便墜落事故  |
| 4        | エボラ出血熱           | 地方病_(日本住血吸虫症)   |
| 4        | エボラ出血熱           | 避難勧告            |
| 4        | エボラ出血熱           | マールブルグ熱         |
| 4        | エボラ出血熱           | 重症急性呼吸器症候群      |
| 4        | エボラ出血熱           | シエラレオネ          |
| 4        | エボラ出血熱           | リベリア            |
| 5        | 乃木坂 46           | 秋元康             |
| 5        | 乃木坂 46           | 前田敦子            |
| 5        | 乃木坂 46           | 大島優子            |
| 5        | 白石麻衣             | 生駒里奈            |
| 6        | ドラゴンクエスト         | ファイナルファンタジー X   |
| 6        | ドラゴンクエスト         | ファイナルファンタジー XI  |
| 6        | ドラゴンクエスト         | ファイナルファンタジー X-2 |
| 6        | ドラゴンクエスト         | 無双シリーズ          |
| 8        | 俺の妹がこんなに可愛いわけがない | irony           |
| 8        | 俺の妹がこんなに可愛いわけがない | 伏見つかさ           |
| 10       | ラブライブ            | iPhone アプリ      |
| 11       | SWITCH           | ゲーム機            |
| 11       | 榮倉奈々             | 佐藤健             |

ID:7 と ID:9 以外の実験協力者は関係のあるアイテムを発見できていた。ID:7 の実験協力者は LiSa という歌手に関係するアイテムを発見しようとしたが、LISA という同じ名前の異なる歌手に対して検索を行っていたため発見できなかったと考える。ID:9 の実験協力者は歌手のゴールデンボンバーやヨイトマケの唄などに関係するアイテムを発見しようとしたが、論理演算機能を利用

しておらず、適切な絞り込みを行わなかったため発見できなかったと考える。また、ID:8 はアイテムに関係のあるアイテムを発見するのではなく、自分がアイテムとの関係を知っているアイテムの動向情報を確認し、関係を確認するという作業を行っていた。これは、タスクの説明不足による外れ値だと考える。

ID:8 の発見したアイテムを除き、発見したアイテム 26 のうち 10 アイテムにおいて、アイテム発見時に論理演算機能を利用していなかった。9 アイテムが 2 回の検索結果に対して論理演算機能を用いて、3 アイテムが 3 回の論理演算機能を用いて、4 アイテムが 4 回の検索結果に対して論理演算機能を用いて発見しており、予備実験に比べて、複雑な検索が行われていることがわかった。

実験協力者がタスク 2 において何ページ目でアイテムを発見したかを図 31 で示す。

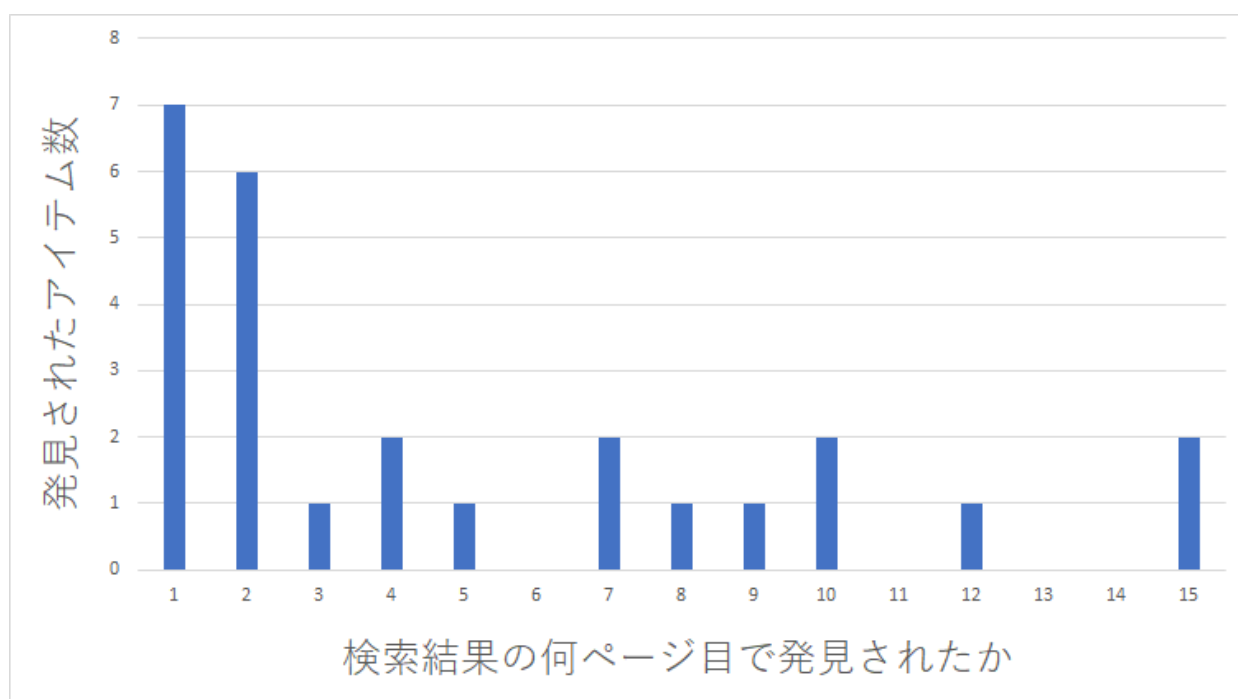


図 31 実験協力者がアイテムを発見したページ

図 31 を確認すると、最初の 1 ページ目で発見されたアイテムが 1 番多かった。また、5 ページ目以降でアイテムを発見している実験協力者は ID:2, ID:4, ID:6 の 3 名であった。予備実験やタスク 1 と比較してランキング下位で発見されるアイテムが多かった。これは、検索対象が自分の興味・関心のあるアイテムであったのでランキング下位まで確認したこと要因と考える。

実験協力者の特徴的変動の利用回数を表 20、論理演算子の利用回数を表 21 で示す。

表 20 特徴的変動の利用回数

| 実験協力者 ID | ALL | MAX | MIN | SI | SD | PEAK | BOTTOM | NONE |
|----------|-----|-----|-----|----|----|------|--------|------|
| 1        | 1   | 0   | 0   | 0  | 0  | 28   | 0      | 0    |
| 2        | 10  | 17  | 0   | 22 | 0  | 21   | 0      | 0    |
| 3        | 10  | 5   | 0   | 2  | 0  | 19   | 0      | 0    |
| 4        | 2   | 0   | 0   | 10 | 0  | 11   | 0      | 0    |
| 5        | 3   | 1   | 0   | 8  | 0  | 48   | 0      | 0    |
| 6        | 1   | 0   | 0   | 0  | 0  | 16   | 0      | 0    |
| 7        | 4   | 2   | 0   | 0  | 0  | 17   | 0      | 0    |
| 8        | 6   | 0   | 0   | 0  | 0  | 6    | 0      | 0    |
| 9        | 10  | 10  | 0   | 0  | 0  | 7    | 0      | 0    |
| 10       | 23  | 0   | 0   | 2  | 0  | 11   | 0      | 0    |
| 11       | 6   | 0   | 0   | 26 | 0  | 0    | 0      | 0    |
| 計        | 76  | 35  | 0   | 70 | 0  | 184  | 0      | 0    |

表 21 論理演算子の利用回数

| 実験協力者 ID | 利用なし | AND | OR | NOT |
|----------|------|-----|----|-----|
| 1        | 24   | 24  | 1  | 0   |
| 2        | 53   | 14  | 5  | 0   |
| 3        | 22   | 15  | 0  | 0   |
| 4        | 18   | 5   | 0  | 0   |
| 5        | 42   | 18  | 0  | 0   |
| 6        | 11   | 6   | 0  | 0   |
| 7        | 14   | 9   | 0  | 0   |
| 8        | 12   | 0   | 0  | 0   |
| 9        | 27   | 0   | 0  | 0   |
| 10       | 27   | 9   | 0  | 0   |
| 11       | 19   | 13  | 0  | 0   |
| 計        | 269  | 113 | 6  | 0   |

特徴的変動は PEAK, ALL, SI, の順に利用回数が多かった。予備実験では, MAX が 1 番利用されており ALL はほとんど利用されていなかった。このような, 特徴的変動の利用頻度の変化は効率的な検索行動に対する知識を提示したことが要因と考える。4.2 節で述べた, 特徴的変動 MAX, MIN と論理演算機能の組み合わせ方で提供した知識によって MAX の利用頻度が低くなっ

たを考える。また ALL は、3.3 節で述べた情報要求を適切に表現するための主題分析で提供した知識によって、実験協力者がアイテムについての動向情報を確認するために利用頻度が高くなったと考える。

論理演算機能において、NOT は利用されなかったが、他の論理演算子の利用頻度に変化は見られなかった。また、ID:8、ID:9 の実験協力者は論理演算機能を利用しなかった。ID:8 の実験協力者は、検索結果のタイプ@period しか利用しておらず、適切な検索を行えていなかった。ID:9 の実験協力者は、予備実験の時点で論理演算機能をほとんど利用していなかった。今回の実験協力者には予備実験も行ってもらっているため、改めてコンテキスト検索エンジンのガイドラインを確認してもらうことはしなかった。予備実験から 1 月ほど期間が空いてしまったので、予備実験の段階で論理演算機能に対する理解が不足していたと考えられる ID:9 の実験実験者だけでなく、ID:8 の実験協力者も論理演算機能に対する理解が低下したと考える。

実験後アンケートの回答結果を図 32、図 33 で示す。図 32 はフィードバックされた効率的な検索行動に対する知識のわかりやすさを 5 段階で評価してもらった。図 33 はフィードバックされた知識によって検索行動に変化があったかを 5 段階で評価してもらった。

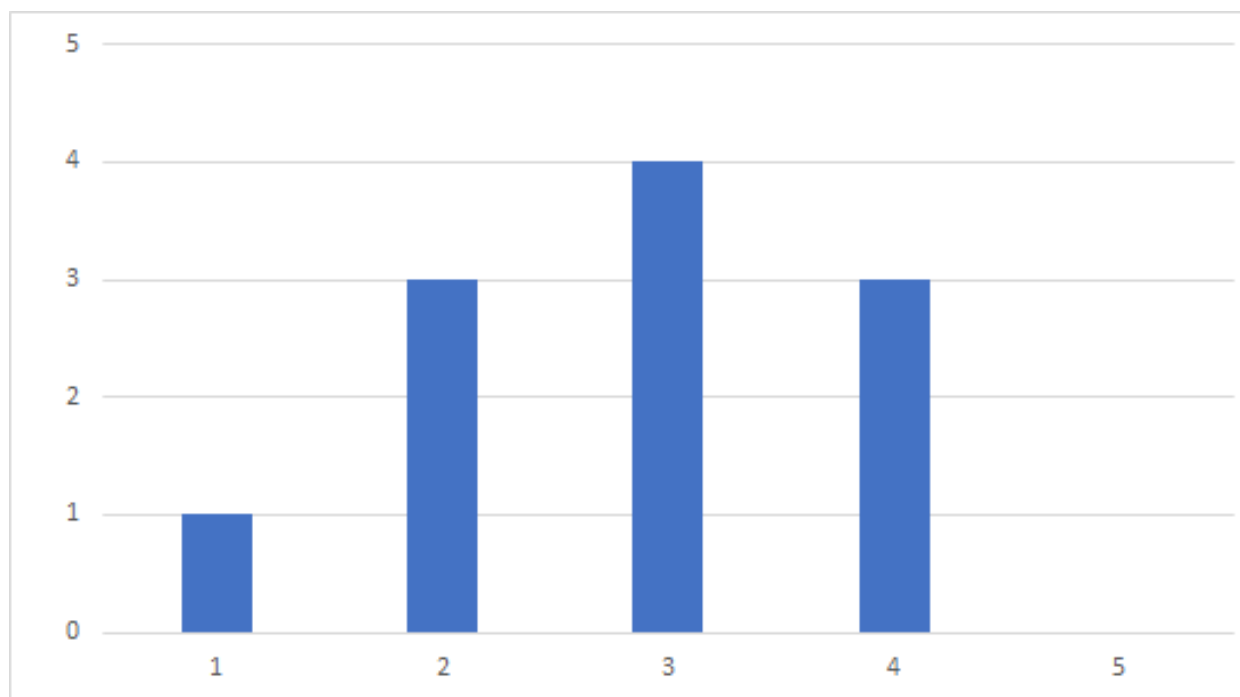


図 32 効率的な検索行動に対する知識のわかりやすかったか？ (1:わかりやすい～5:わかりにくい)

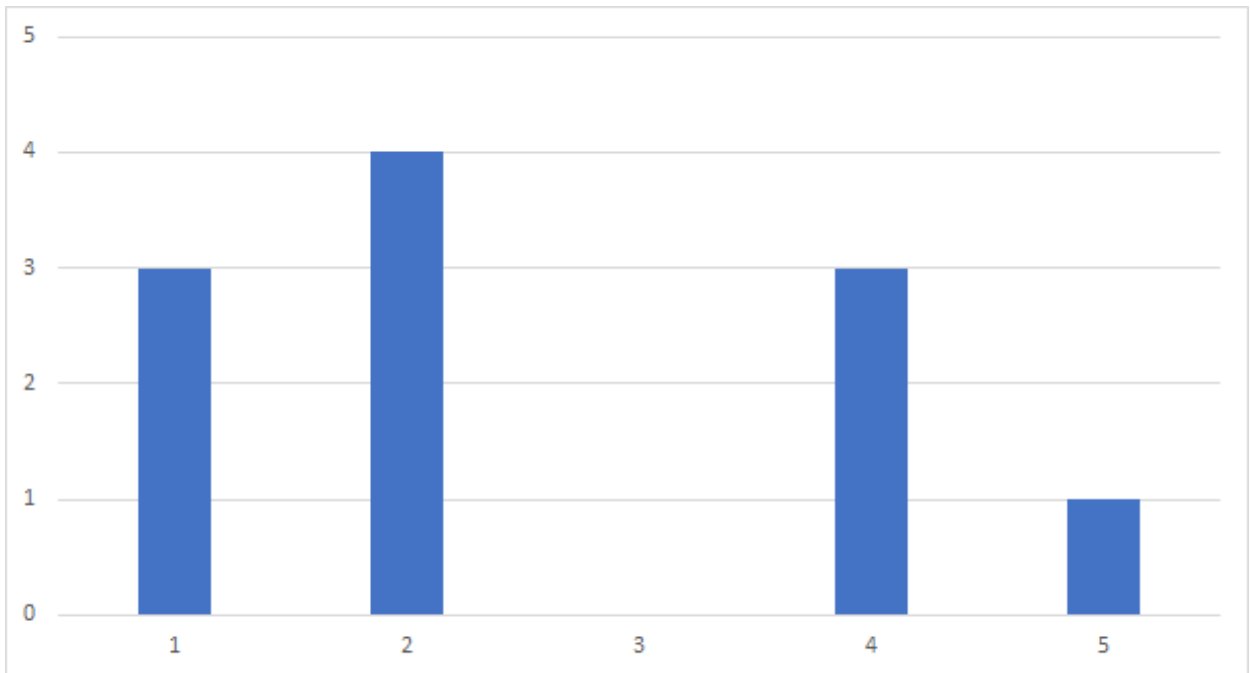


図 33 フィードバックされた知識によって検索行動に変化はあったか？（1:変化があった～5:変化がなかった）

図 32, 図 33 を確認すると, 提供した知識は実験協力者によって大きく評価が分かれた. 効率的な検索行動に対する知識の理解度と行動の変化について, ID:3 の実験協力者は理解度・行動の変化共に 1 と高評価であったが, 同じ理解度 3 の評価をしたが, 行動の変化の評価は 1 と 5 に分かれたりと, 実験協力者によって評価が分かれた. また, ID:6 の実験協力者は理解度を 2, 行動の変化を 4 と評価し, 検索行動に変化はほとんど変化がなかったと回答したが, 提供した知識に基づき網羅的なアイテムの確認を行っており, それによって, 予備実験ではアイテムを 1 つしか発見できていなかったが, 本実験では両タスク合計で 14 アイテムも発見できていた. 検索行動に変化があったと回答した実験協力者の検索行動の変化としては, クエリ生成で絞り込みを意識した, 検索結果を 1000 件以下まで絞り込んでからアイテムを確認した, アイテム名の部分一致検索を用いて検索するアイテム名の妥当性を検証した等が挙げられる. 本実験で事前に記事又は自分の好きなものとの関連に気づいていなかった, または意外なアイテムを発見できたと回答した実験協力者は 7 名であり予備実験と同じであった. ID:5, ID:8 の実験協力者は, 予備実験では関係に気づいていなかったアイテム, または意外なアイテムを発見できていたが, 今回の実験では発見できなかった. ID:5 の実験協力者に関しては, 感想に元々疲れがたまっており, 実験に集中ができていなかったとの回答があり, これが原因であると考え. ID:8 の実験協力者に関しては, 前述したとおり, 実験の期間が空いたことによるコンテキスト検索エンジンに対する基本知識の低下が原因であると考え. 本実験のみで関係に気づいていなかった, または意外なアイテムを発見できた ID:6 と ID:7 は, 効率的な検索行動に対する知識を提供したことによる, 検索結果の効率的な絞り込みや網羅的なアイ

テム確認によるものと考える。

以上の結果から、効率的な検索行動に対する知識を提供することによる、ユーザの効率的な検索行動支援は個人差は大きいが有効性があると考える。

## 5 おわりに

本論文では、コンテキスト検索エンジンへの論理演算機能の導入を提案した他、ガイドラインの作成、効率的な検索行動の提示を行うことでコンテキスト検索エンジンの効率的な活用支援を提案した。

提案手法の有効性を検証するため、4つの検索タスクを論理演算機能を導入したコンテキスト検索エンジンと従来のコンテキスト検索エンジンで行う比較実験を行った。その結果、年単位で周期的なアイテムや時期の異なる同一イベントに関連するアイテムの発見に対して有効性があることを示した。また、ユーザの興味によって検索タスクを決定する、従来の Web 検索エンジンの利用に近いタスクを想定した予備実験を行い、収集した検索行動のログを分析した結果から効率的な検索行動に対する知識を作成した。この知識を実験協力者に提示し行った本実験では、複雑な検索クエリを利用した検索や、検索結果の網羅的な確認などの検索行動が確認されており、作成したガイドラインや効率的な検索行動に関する知識の有効性を示した。

今後、Web 上の動向情報の増加とユーザの情報要求の多様化に伴い、コンテキスト検索エンジンの活用できる場面が広がっていくと考える。その際、本研究の成果はコンテキスト検索エンジンの効率的な活用に貢献することが期待できる。



本研究は、首都大学東京大学院システムデザイン研究科システムデザイン選考情報通信システム学域において、多くの方々のご指導、ご協力のもとに進めたものであります。指導教員である高間康史教授には本研究に関する御指導、御助言、御助力を賜りました。心より感謝いたします。また、本研究を進めるにあたり、貴重な御指導、御助言をいただいた、山口亨教授、片山薫准教授、小町守准教授にも深く感謝いたします。

日頃より研究に関する相談や助言、実験に協力いただきました、高間研究室の皆様にも感謝いたします。

## 参考文献

- [1] 小林 竜己: Web 検索エンジンの技術動向, 特許技術懇談会誌, No. 252, pp. 83-89, 2009.
- [2] A. Spink, D. Wolfarm, M. B. J. Jansen, T. Saracevic: Searching the Web: The Public and Their Queries, Journal of the American Society for Information Science and Technology, Vol. 52, No. 3, pp. 226-234, 2001.
- [3] J. Huang, E. N. Efthimiadis: Analyzing and Evaluating Query Reformulation Strategies in Web Search Logs, Proceedings of the 18th ACM Conference on Information and Knowledge Management, pp. 77-86, 2009.
- [4] 今井 良太, 戸田 浩之, 関口 裕一郎, 望月 崇由, 鈴木 智也, 今井桂子: Web 検索サービスにおける多義的なクエリ推薦手法, 日本データベース学会論文誌, Vol. 9, No. 1, pp. 7-11, 2010.
- [5] 森 辰則, 福本 淳一, 加藤 恒昭, 梶井 文人, 佐々木 裕, H. Chen, K. Chen, C. Lin, 三田村 照子, E. Nyberg, 神門 典子: NTCIR における質問応答技術の評価と今後の展望, 情報処理学会研究報告自然言語処理, Vol. 2008, No. 4, pp. 43-50, 2008.
- [6] 徳永 健伸: 言語処理を利用した知的情報アクセス: 検索, 抽出, 要約, 分類, QA, オペレーションズ・リサーチ: 経営の科学, Vol. 52, No. 11, pp. 713-718, 2007.
- [7] 田村 元秀, 村上 仁一, 徳久 雅人, 池原 悟: Web 検索エンジンを用いた Why 型質問応答システムに関する研究, 情報処理学会研究報告, Vol. 2008, No. 4, pp. 15-21, 2008.
- [8] 三原 英理, 藤井 敦, 石川 徹也: 行動表現に着目したヘルプデスク志向の質問応答, 言語処理学会第 11 回年次大会発表論文集, pp. 1096-1099, 2005.
- [9] A. McCakkum, K. Nigam, J. Rennie, K. Seymore: A Machine Learning Approach to Building Domain-Specific Search Engine, Proceedings of the 16th International Joint Conference on Artificial Intelligence, Vol. 2, pp. 662-667 1999.
- [10] 亀井 俊之, 門田 暁人, 松本 健一: WWW を対象としたソフトウェア検索エンジンの構築, 電子情報通信学会ソフトウェアサイエンス研究会, Vol. 102, No. 617, pp. 59-64, 2003.
- [11] J. Shakes, M. Langheinrich, O. Etzioni: Dynamic Reference Sifting: a Case Study in the Homepage Domain, Proceedings of the 6th International World Wide Web Conference, pp. 189-200, 1997.
- [12] 小久保 卓, 小山 聡, 山田 晃弘, 北村 泰彦, 石田 亨: 検索隠し味を用いた専門検索エンジンの構築, 情報処理学会論文誌, Vol. 43, No. 6, pp 1804-1813, 2002.
- [13] S. Oyama: Domain-Specific Web Search with Keyword Spices, IEEE Transactions on Knowledge and Data Engineering, Vol. 6, No. 1, pp. 17-27, 2004.
- [14] 高間 康史, 加藤 優, 桑折 章吾, 石川 博: 動向に関する問いを対象とした検索エンジンの提案, 人工知能学会論文誌, Vol. 30, No. 1, pp. 138-147, 2015.
- [15] 高間 康史, Y. Zhu, 桑折 章吾, 山口 晃一, 瀧口 慈勇: 動向に関する問いに答えるコンテキスト検索エンジンの開発, 人工知能学会第 10 回インタラクティブ情報アクセスと可視化マイニ

- ング研究会, pp. 9-15, 2015.
- [16] 桑折 章吾, 加藤 優, 高間 康史: 検索エンジンを用いた情報検索におけるユーザ行動の分析, 人工知能学会第4回インタラクティブ情報アクセスと可視化マイニング研究会, pp. 9-14, 2013.
  - [17] 加藤 恒昭, 松下 光範, 平尾 努: 動向情報の要約と可視化に関するワークショップの提案, 情報処理学会自然言語処理研究会, Vol. 2004, No. 108, pp. 89-94, 2004.
  - [18] M. Yoshida, Y. Arase, T. Tsunoda, M. Yamamoto: Wikipedia Page View Reflects Web Search Trend, Proceedings of the ACM Web Science Conference, pp. 65:1-65, 2015.
  - [19] 手塚 拓哉, 山口 晃一, 諸 俊, 桑折 章吾, 高間 康史: コンテキスト検索エンジンへのランキング機能の導入に関する検討, 人工知能学会第11回インタラクティブ情報アクセスと可視化マイニング研究会, pp. 7-11, 2015.
  - [20] Y. Takama, Y. Zhu, S. Kori, K. Yamaguchi, L. H. Chen, H. Ishikawa: Introduction of Sparkline as Snippet to Context Search Engine Result Page, 2017 Conference on Technologies and Applications of Artificial Intelligence, pp. 104-107, 2017.
  - [21] Y. Takama, Y. Zhu, S. Kori, K. Yamaguchi, L. Chen, H. Ishikawa: Design of Context Search Engine Based on Analysis of User's Search Intentions, Journal of Advanced Computational Intelligence and intelligent Informatics, Vol. 20, No. 6, pp. 910-918, 2016.
  - [22] 久野 高志, 安形 輝, 上田 修一: 情報検索システムとしてみたサーチエンジン, 第49回日本図書館情報学会研究大会発表要綱, pp. 337-347, 2001.
  - [23] 原田 昌紀, 佐藤 進也, 風間 一洋: 索引篩法-大規模サーチエンジンのための高速なランキング検索法, 第14回データ工学ワークショップ論文集, 8 pages, 2003.
  - [24] S. Brin, L. Page: The Anatomy of a Large-Scale Hypertextual Web Search Engine, Computer Networks and ISDN System, Vol. 30, No. 1-7, pp. 107-117, 1998.
  - [25] J. M. Kleinberg: Authoritative Sources in a Hyperlinked Environment, Journal of the ACM, Vol. 46, No. 5, pp. 604-632, 1999.
  - [26] 伊東 民雄: ウェブ情報の検索: 情報源の効率的な探索, 情報管理, Vol. 44, No. 9, pp. 629-640, 2001.
  - [27] 菅谷 狩野: 決定木を用いた情報検索能力の分析, コンピュータ&エデュケーション, Vol. 18, pp. 145-151, 2005.
  - [28] 南 友紀子, 岩瀬 梓, 宮田 洋輔, 石田 栄美, 上田 修一, 倉田 敬子: ウェブ環境における情報検索スキル, 日本図書館情報学会誌, Vol. 60, No. 3, pp. 163-180, 2016.
  - [29] B. J. Jansen, U. Pooch: A Review of Web Searching Studies and a Framework for Future Research, Journal of the American Society for Information Science and Technology, Vol. 52, No. 3, pp. 17-26, 2002.
  - [30] 種市 敦子, 逸村 裕: エンドユーザーの web 探索行動\_短期大学生の実験調査に基づく情報評価モデルの構築, Library and Information Science, No. 55, pp. 1-23, 2006.
  - [31] 齋藤, 三輪: Web 情報検索におけるリフレクションの支援, 人工知能学会誌, Vol. 19, No. 4, pp. 214-224, 2004.

- [32] 原田 智子: サーチャーに必要な基礎知識, 情報の科学と技術, Vol. 38, No. 8, pp. 433-439, 1988.
- [33] 岡 紀子: データベースの検索からみた情報専門家の役割と将来像, 情報の科学と技術, Vol. 67, No. 9, pp. 465-471, 2017.
- [34] 田中 早苗: 医薬品安全情報のオンラインデータベースでの検索テクニック, 情報の科学と技術, Vol. 57, No. 6, pp. 292-297, 2007.
- [35] 石田 由利子: エンドユーザー教育と検索システムの選択, 情報の科学と技術, Vol. 54, No. 5, pp. 240-247, 2004.
- [36] 小嶋 智美: EBM を意識した Pubmed の検索 : JMLA 診療ガイドラインワーキンググループの活動, 情報の科学と技術, Vol. 63, No. 5, pp. 186-192, 2013.
- [37] 上田 博三: 新型インフルエンザ対策の経緯, 日本公衆衛生雑誌, Vol. 57, No. 3, pp. 157-164, 2010.

- 佐藤 宏貴, 高間 康史: コンテキスト検索エンジンを利用した時系列データマイニングの提案, JSAI2017, 2M2-OS-34a-3, 2017

## 付録 A コンテキスト検索エンジンのガイドライン

3.2 節で作成し、4.2 節の予備実験で用いたガイドラインを以下に示す。

### コンテキスト検索エンジンガイドライン

コンテキスト検索エンジンには 3 つの基本検索機能が実装されています。

現在利用可能な入力フォームには「期間」「アイテム名」「単位」「データベース」「特徴的変動」「(特徴的変動が”UP”, ”DOWN”の時) 数字 or 平均値」「検索結果」「論理演算子」があります。

図 1. コンテキスト検索エンジンの入力フォーム

特徴的変動は現在、以下の 9 つが用意されています。

| 特徴的変動          | グラフ概形 | 抽出条件                               |
|----------------|-------|------------------------------------|
| 最大値<br>(MAX)   |       | 各動向情報が最大の<br>期間                    |
| 最小値<br>(MIN)   |       | 各動向情報が最少の<br>期間                    |
| 急上昇<br>(SI)    |       | 3 ヶ月以内に、変動幅<br>の 20%以上の増加          |
| 急下降<br>(SD)    |       | 3 ヶ月以内に、変動幅<br>の 20%以上の減少          |
| ピーク<br>(PEAK)  |       | 変動幅の 10%以上の<br>増加後、減少に転じた<br>山の頂点  |
| 底<br>(BOTTOM)  |       | 変動幅の 10%以上の<br>減少後、増加に転じた<br>谷の頂点  |
| 安定<br>(NONE)   |       | 変動幅が月に 1.5%を<br>超えない 3 ヶ月以上<br>の期間 |
| 基準値以上<br>(UP)  |       | 指定した基準値以上<br>の値をとる期間               |
| 基準値以下<br>(LOW) |       | 指定した基準値以下<br>の値をとる期間               |

図 2. 特徴的変動の説明

これらの特徴的変動に、すべての特徴的変動を検索する「ALL」を加えた 10 個の特徴的変動を利用することができます。（本実験では、「UP」「DOWN」は利用しないので、これらの利用例にも利用しません。）

これから、3 つの基本検索機能の利用方法について説明していきます。

#### 1. 指定したアイテムに関する動向が特徴的変動を示した期間の検索

これは検索したいアイテムが分かっている際に、そのアイテムの特徴的変動がいつ発生したかを検索する機能です。

使用例 1: 「ポケットモンスター」のアイテムがいつ特徴的変動「PEAK」を迎えたかを知りたい！

検索クエリ 1: 「アイテム名」に「ポケットモンスター」、「Variation:ALL」（特徴的変動）を「PEAK」、「@item」（検索結果）を「@period」にし、「Search」をクリック

The screenshot shows a search interface with the following fields and values:

- StartYear: [dropdown]
- StartMonth: [dropdown]
- EndYear: [dropdown]
- EndMonth: [dropdown]
- 完全一致: [dropdown]
- 検索対象: ポケットモンスター
- Unit: [dropdown]
- All source: [dropdown]
- Variation: PEAK
- input: input AVG or any number
- @Period: [dropdown]
- Target: NONE
- Search Button: Q Search

図 3. 検索クエリ 1 の入力フォーム

“ポケットモンスター PEAK @period” の検索結果 ( 5 )

[クエリを修正する](#)

| # | Item      | Unit | DBsource           | VarType | Period            | Value & Sparkline | Target | Google |
|---|-----------|------|--------------------|---------|-------------------|-------------------|--------|--------|
| 1 | ポケットモンスター | ビュ   | Wikipedia pageview | PEAK    | 2008-08 ~ 2008-10 | 101744.0<br>      | Google |        |
| 2 | ポケットモンスター | ビュ   | Wikipedia pageview | PEAK    | 2009-04 ~ 2009-06 | 79429.0<br>       | Google |        |
| 3 | ポケットモンスター | ビュ   | Wikipedia pageview | PEAK    | 2010-08 ~ 2010-10 | 246569.0<br>      | Google |        |

図 4. 検索クエリ 1 の検索結果(上位 3 件)

このようなクエリで検索すると、図 4 のような結果が表示されます。  
 検索クエリとその検索クエリでの検索結果の数の下に検索結果が表示されます。

このクエリで検索を行うとアイテム名は完全一致したもののみを返します。なので、以下のような例には正しい検索結果が得られない場合があります。

例 1. 「ポケモン」と「ポケットモンスター」では検索結果が違う。

例 2. 「少女時代」で検索しても検索結果が少ない。

➡ Wikipedia の記事名が「少女時代\_(音楽グループ)」だった。

アイテム名の左隣にある「完全一致」を「部分一致」にすることでクエリのアイテム名で部分一致しているアイテムを検索することができます。

使用例 2: アイテム名に「嵐」が含まれているアイテムがいつ特徴的変動「MAX」を迎えたかを知りたい！

検索クエリ 2: 「完全一致」を「部分一致」、「アイテム名」に「嵐」、「Variation:ALL」  
 (特徴的変動)を「MAX」、「@item」(検索結果)を「@period」にし、「Search」  
 をクリック

図 5. 検索クエリ 2 の入力フォーム

“嵐(部分一致) MAX @period” の検索結果 (626)  
 ▼クエリの修正

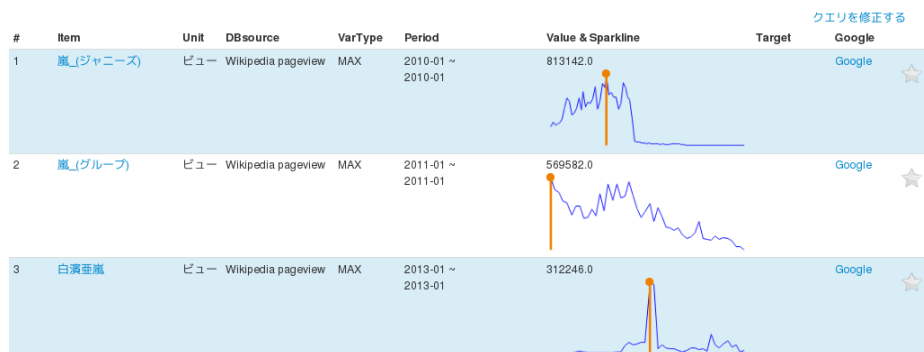


図 6. 検索クエリ 2 の検索結果(上位 3 件)



※ 部分一致検索は、検索したいアイテムの正しい表記を発見するのに有効な手段なので、特徴的変動は各アイテムに 1 つしかない「MAX」・「MIN」を選択した方がよいです。そして、正しい表記がわかったら、そのアイテム名で検索した方がよいです。

## 2. 指定した期間に特徴的変動を示したアイテム(動向)の検索

これは検索したい期間が分かっている際に、その期間に特徴的変動が発生しているアイテムを検索する機能です。

使用例 3: 2013 年 1 月～6 月に特徴的変動「SI」が発生しているアイテムを知りたい！

検索クエリ 3: 「StartYear」に「2013 年」、「StartMonth」に「1 月」、「EndYear」に「2013 年」、「EndMonth」に「6 月」、「Variation:ALL」(特徴的変動)を「SI」、  
「@period」(検索結果)は「@item」にし、「Search」をクリック

図 7. 検索クエリ 3 の入力フォーム

このようなクエリで検索すると、図 8 のような結果が表示されます。

検索結果を「@item」(アイテム)にすると、アイテムごとにスコア付けされており、検索結果はスコアが高い順にソートされています(ランキング機能)。

“2013/1-2013/6 SI @item” の検索結果 (668910)

| # | Item  | Unit | DBsource              | VarType | Period               | Value & Sparkline | Target | Google |
|---|-------|------|-----------------------|---------|----------------------|-------------------|--------|--------|
| 1 | 進撃の巨人 | ビュー  | Wikipedia<br>pageview | SI      | 2013-04 ~<br>2013-05 | 479150.0<br>      | Google | ★      |
| 2 | 進撃の巨人 | ビュー  | Wikipedia<br>pageview | SI      | 2013-02 ~<br>2013-04 | 1142000.0<br>     | Google | ★      |
| 3 | あまちゃん | ビュー  | Wikipedia<br>pageview | SI      | 2013-02 ~<br>2013-04 | 533590.0<br>      | Google | ★      |

図 8. 検索クエリ 3 の検索結果(上位 3 件)

※特徴的変動「MAX」「MIN」はアイテムごとに一番高い(低い)期間である 1 点を示しています。しかし、特徴的変動「SI」「SD」では 3 ヶ月以内に 20%以上増加(減少)した時の始点と終点の 2 点の期間を示しており、特徴的変動「PEAK」「BOTTOM」の場合は増加(減少)した後に減少(増加)した時の頂点を示しており、山(谷)の始点・頂点・終点の 3 点の期間が必要になります。

コンテキスト検索エンジン内のデータのほとんどが月別なので、期間を指定する検索を行う場合、特徴的変動に合わせた期間を選択する必要があります。下記に例を示します。

例 1. 2011 年 3 月に「MAX」 ➡ 期間指定「2011 年 3 月～2011 年 3 月」

例 2. 2011 年 3 月に「SI」 ➡ 期間指定「2011 年 2 月～2011 年 3 月」

例 3. 2011 年 3 月に「PEAK」 ➡ 期間指定「2011 年 2 月～2011 年 4 月」

### 3. 指定したアイテムに関する動向が特徴的変動を示した期間に同様又は逆の変動を示したアイテム(動向)の検索

この検索方法は基本検索機能 1 と基本検索機能 2 の 2 つを組み合わせたような検索です。まずは、入力されたアイテム名を用いて、そのアイテムの特徴的変動が発生した時期を検索します。そして、結果として得られた複数の期間の中で同じ又は逆の特徴的変動が発生しているアイテムを結果として返します。

使用例 4: 「ポケットモンスター」のアイテムに特徴的変動「PEAK」が発生した期間に特徴的変動「BOTTOM」が発生しているアイテムを知りたい！

検索クエリ 4: 「アイテム名」に「ポケットモンスター」、「Variation:ALL」(特徴的変動)を「PEAK」、「@item」(検索結果)を「@item opposite」にし「Search」をクリック

図 9. 検索クエリ 3 の入力フォーム

このようなクエリで検索すると、図 8 のような結果が表示されます。

検索結果を「@item oppsite」(アイテム)にすると、検索結果「@item」と同様にスコアが高い順にソートされています(ランキング機能)。

“ポケットモンスター PEAK OPPOSITE @item” の検索結果 ( 493019 )

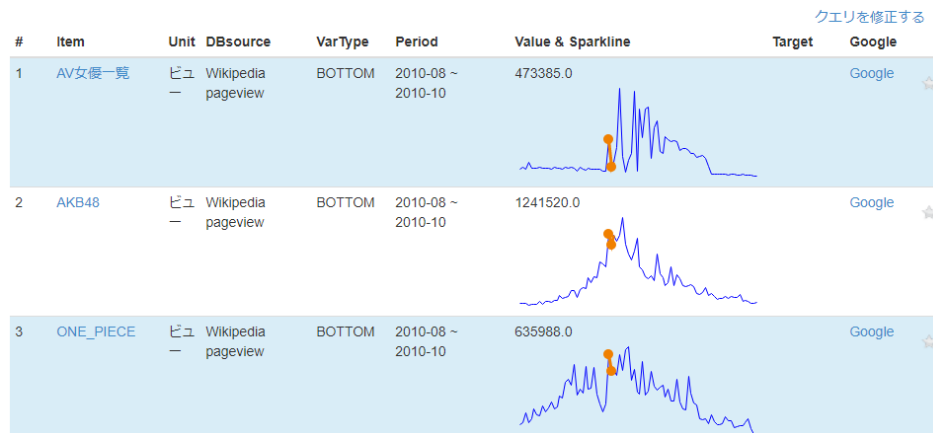


図 10. 検索クエリ 2 の検索結果(上位 3 件)

検索結果のアイテム名をクリックすることで、アイテムページに飛び、アイテムの動向情報詳細を見ることが出来ます。今の検索結果から 2 番目にある AKB48 のアイテム名をクリックします。

#### AKB48のWikipedia pageviewについてのグラフ

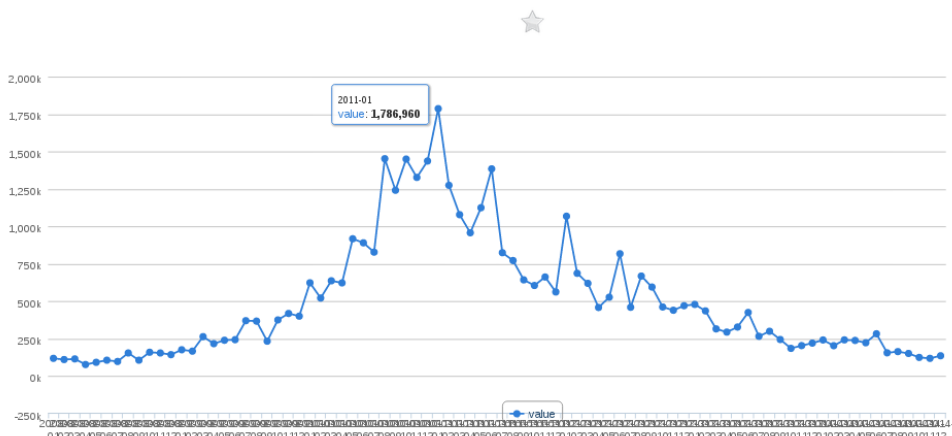


図 11. AKB48 のアイテムページ

アイテムページでは、マウスを点に合わせると時期とその時の値を知ることが出来ます、図 11 の例では AKB48 の Wikipedia は 2011 年 1 月に 1,786,960 回閲覧されたことが分かります。

上記の3つの基本検索機能とアイテムページに加えて、検索結果の絞り込みのために「Unit」(単位)と「source」(データベース)を選択することができます。「Unit」は「円」「件」「ポイント」「ビュー」の様にアイテムの単位を指定して検索ができます。「source」では Wikipedia のページビューのデータのみを利用した「only WPV」、Wikipedia のページビューを除いたデータのみを利用した「exclude WPV」、どちらのデータも利用した「ALL source」が選択できます。

## コンテキスト検索エンジンの論理演算

コンテキスト検索エンジンには新たに検索結果同士を論理演算子でつなぐことで絞り込みを行う機能があります。論理演算子は「AND」「OR」「NOT」の3つがあります。これらの論理演算子の利用方法を説明していきます。

### 1. 論理演算子「AND」

論理演算子「AND」を利用することで、2つの検索結果の両方に含まれているものだけを検索結果として得ることができます。

使用例 1: 2012 年 3 月に「PEAK」が発生し、かつ 2012 年 9 月に「PEAK」が発生しているアイテムを知りたい！

検索フロー1:

- ① 「StartYear」に「2012 年」、「StartMonth」に「2 月」、「EndYear」に「2012 年」、「EndMonth」に「4 月」、「Variation:ALL」(特徴的変動)を「PEAK」、「@item」(検索結果)は「@item」にし、「Search」をクリック
- ② ①の検索結果が表示されたのを確認してから、次の検索クエリを入力
- ③ 「StartYear」に「2012 年」、「StartMonth」に「8 月」、「EndYear」に「2012 年」、「EndMonth」に「10 月」、「Variation:ALL」(特徴的変動)を「PEAK」、「@item」(検索結果)は「@item」にし、「NONE」(論理演算子)を「@AND」にし、「Search」をクリック(図 12)

The screenshot shows a search interface with the following elements:

- Top row: Date range selectors for Start (2012年 8月) and End (2012年 10月).
- Second row: Search criteria dropdowns: 完全一致, item (e.g. 乾電池, にんじん, 自転車), Unit, and All source.
- Third row: Logical operator dropdowns: PEAK, input AVG or any number, @item, and @AND.
- Bottom row: A blue button labeled "Search".

図 12, 検索フロー1 の③の入力フォーム

このような検索フローで検索を行った結果が図 13 になります。

この検索結果は①の検索クエリと③の検索クエリのどちらにも当てはまる検索結果のみが表示されています。このように複数の検索クエリを「AND」でつなぐことによって検索結果を絞り込むことができます。各クエリの隣の数字がそのクエリ時点での検索件数になっています。よってこの例では、1つ目のクエリの検索結果が 233870 件、2つ目のクエリの隣の数字は1つ目のクエリと 2つ目のクエリの AND 検索した検索結果の件数であり、この例では 44194 件になります。

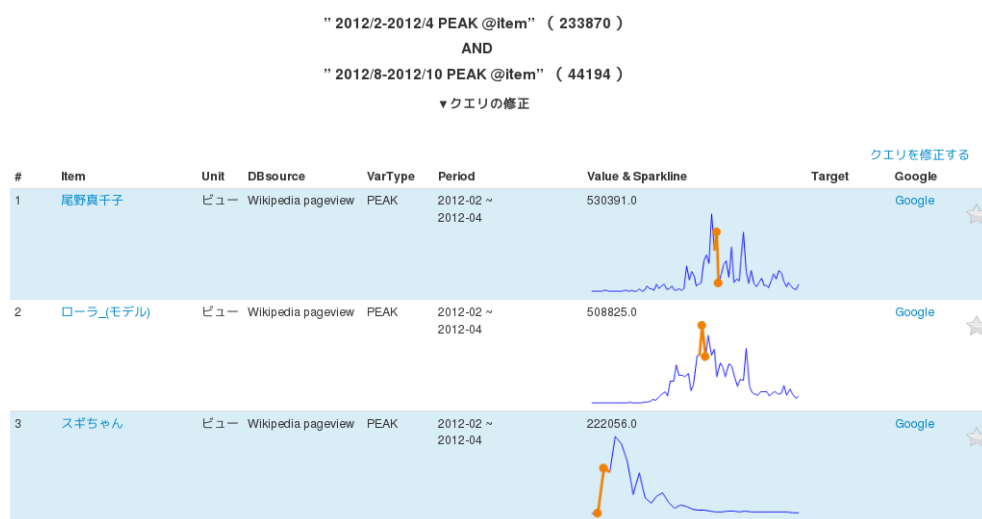


図 13. 検索フロー1 の検索結果(上位 3 件)

## 2. 論理演算子「OR」

論理演算子「OR」を利用することで、2つの検索結果のどちらか一方にでも含まれているものを検索結果として得ることができます。

使用例 2: アイテム「ポケットモンスター」の「PEAK」が発生している期間、またはアイテム「ポケモン」の「PEAK」が発生している期間を知りたい！

検索フロー2:

- ① 「アイテム名」に「ポケットモンスター」、「Variation:ALL」(特徴的変動)を「PEAK」、「@item」(検索結果)を「@period」、「@AND」(論理演算子)を「NONE」にし「Search」をクリック
- ② ①の検索結果が表示されたのを確認してから、次の検索クエリを入力
- ③ 「アイテム名」に「ポケモン」、「Variation:ALL」(特徴的変動)を「PEAK」、「@item」(検索結果)を「@period」、「NONE」(論理演算子)を「@OR」にし、「Search」をクリック(図 14)

図 14. 検索フロー2 の③の入力フォーム

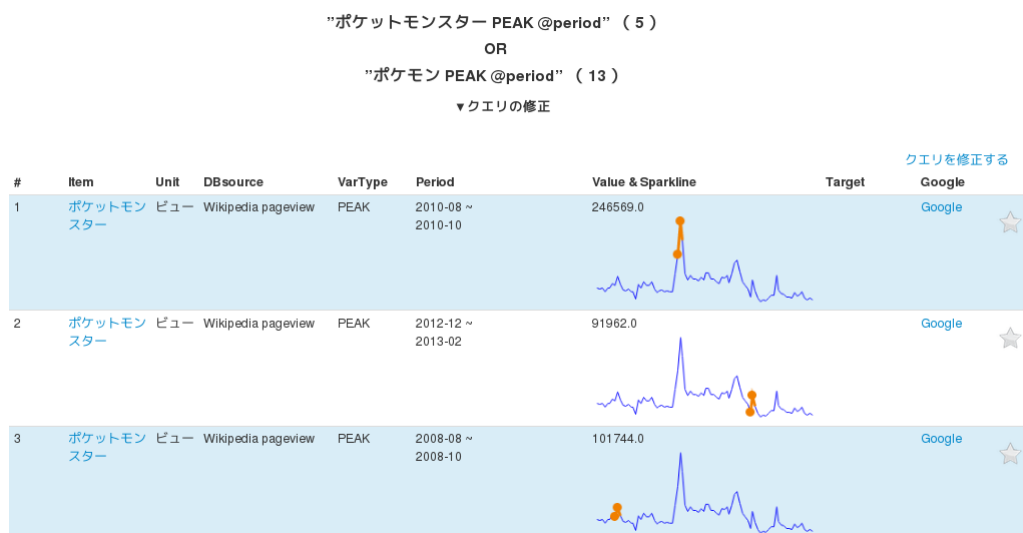


図 15. 検索フロー2 の検索結果(上位 3 件)

このような検索フローで検索を行った結果が図 15 になります。

この検索結果は①の検索クエリと③の検索クエリのどちらか一方でも当てはまる検索結果が表示されています。このように複数の検索クエリを「OR」でつなぐことによって検索結果を広げることができます。

### 3. 論理演算子「NOT」

論理演算子「NOT」を利用することで、1 つめの検索結果から 2 つめの検索結果に含まれているものを取り除くことができます。

使用例 3: 2010 年 3 月に「PEAK」が発生しているアイテムの中で、2009 年 3 月に「PEAK」が発生していないアイテムを知りたい！

検索フロー3:

- ① 「StartYear」に「2010 年」、「StartMonth」に「2 月」、「EndYear」に「2010 年」、「EndMonth」に「4 月」、「Variation:ALL」(特徴的変動)を「PEAK」、「@item」(検索結果)は「@item」、「@OR」(論理演算子)を「START/RESTART」にし「Search」をクリック
- ② ①の検索結果が表示されたのを確認してから、次の検索クエリを入力
- ③ 「StartYear」に「2009 年」、「StartMonth」に「2 月」、「EndYear」に「2009 年」、「EndMonth」に「4 月」、「Variation:ALL」(特徴的変動)を「PEAK」、「@item」(検索結果)は「@item」にし、「START/RESTART」(論理演算子)を「@NOT」にし、「Search」をクリック(図 16)

図 16. 検索フロー3 の③の入力フォーム

このような検索フローで検索を行った結果が図 17 になります。

この検索結果は①の検索クエリの検索結果の中から③の検索クエリの検索結果を取り除いた検索結果が表示されています。このように複数の検索クエリを「NOT」でつなぐことによって検索結果を絞り込むことができます。「NOT」は前述の「AND」「OR」と異なり、先の検索結果から後の検索結果を取り除くので順番によって結果が異なります。

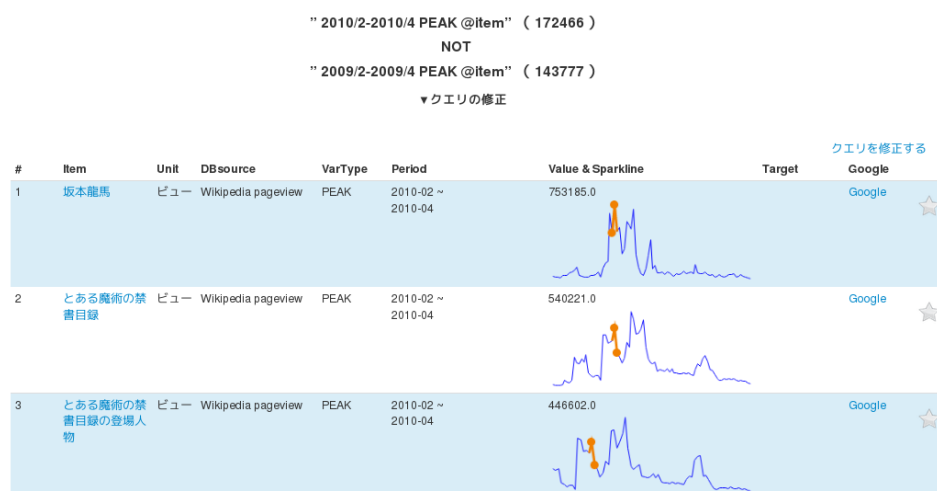


図 17. 検索フロー3 の検索結果(上位 3 件)

検索例では 2 つのクエリの論理演算式のみを対象としましたが、実際はもっと多くのクエリをつなげる事で、検索結果をより絞り込むことができます。

コンテキスト検索エンジンの論理演算は選言標準形を採用しています。

選言標準形とは、論理式の正規化の一種で論理積の論理和の形式になっています。

選言標準形には以下のような例があります。

例1. (A and B) or C  
 例2. (A and notB and notC) or (notD and E and F)  
 論理積: and    否定: not    論理和: or

コンテキスト検索エンジンでの選言標準形の検索例として検索フロー4 を実行する。  
 検索フロー4:

- ① StartYear」に「2012 年」、「StartMonth」に「1 月」、「EndYear」に「2012 年」、「EndMonth」に「3 月」、「Variation:ALL」(特徴的変動)を「MAX」、「@item」(検索結果)は「@item」にし、「@NOT」(論理演算子)を「NONE」にし、「Search」をクリック
- ② StartYear」に「2013 年」、「StartMonth」に「1 月」、「EndYear」に「2013 年」、「EndMonth」に「3 月」、「Variation:ALL」(特徴的変動)を「PEAK」、「@item」(検索結果)は「@item」にし、「NONE」(論理演算子)を「@AND」にし、「Search」をクリック
- ③ StartYear」に「2012 年」、「StartMonth」に「1 月」、「EndYear」に「2012 年」、「EndMonth」に「3 月」、「Variation:ALL」(特徴的変動)を「PEAK」、「@item」(検索結果)は「@item」にし、「@AND」(論理演算子)を「@OR」にし、「Search」をクリック
- ④ StartYear」に「2013 年」、「StartMonth」に「1 月」、「EndYear」に「2013 年」、「EndMonth」に「3 月」、「Variation:ALL」(特徴的変動)を「MAX」、「@item」(検索結果)は「@item」にし、「@OR」(論理演算子)を「@AND」にし、「Search」をクリック

このような検索フローで検索を行った検索クエリが図 18 になります。

```

A " 2012/1-2012/3 MAX @item" ( 70573 )
      AND
B " 2013/1-2013/3 PEAK @item" ( 4593 )
      OR
C " 2012/1-2012/3 PEAK @item" ( 137359 )
      AND
D " 2013/1-2013/3 MAX @item" ( 10172 )
  
```

図 18. 検索フロー4 の検索クエリ



この検索結果を論理式で表すと  $(A \cap B) \cup (C \cap D)$  となります。

この検索では 2012 年 2 月に「MAX」が発生していて、かつ 2013 年 2 月に「PEAK」が発生しているアイテム、又は 2012 年 2 月に「PEAK」が発生していて、かつ 2013 年 2 月に「MAX」が発生しているアイテムを表示しています。このように、複数のクエリをつなげる事で絞り込むことが可能ですが、多くのクエリをつなげることで複雑になり、理解がしづらくなるので注意が必要です。

論理演算を利用する際に、誤ったクエリを入力してしまう可能性があります。このような場合、また 1 から全てのクエリを入力し直すのは必要なく、一部のクエリを修正することが可能です。検索クエリの下の「クエリの修正」をクリックすると、クエリ修正画面が現れるので、

論理演算子を変えるだけの場合は論理演算子を選択し、実行ボタンを押せば再検索ができます。クエリを削除する場合は、削除したいクエリの左の削除ボタンを押すとクエリ修正画面に遷移します。例として、検索フロー4 の検索結果を用いてクエリ修正を行い、3 丁目、4 丁目のクエリを削除し残ったクエリの「NOT」検索を実行しましょう。

検索フロー5：

- ① 検索フロー4 を実行する。
- ② 検索クエリ下の「クエリを修正する」をクリックする。
- ③ 図 19 のページが表示されるので、4 丁目のクエリ 2013/1-2013/3 MAX @item の左にある削除ボタンをクリックする。
- ④ ③の結果、クエリ修正画面に遷移するので、その画面で 3 丁目のクエリ 2012/1-2012/3 PEAK @item の左にある削除ボタンをクリックする。
- ⑤ ④の結果、図 20 のページが表示されるので 2 丁目のクエリ間のラジオボタンを「AND」から「NOT」に変更し、実行ボタンをクリックする。

▼クエリの修正

|  |   |
|--|---|
| <div>削除 2012/1-2012/3 MAX @item</div> <div><input checked="" type="radio"/> AND <input type="radio"/> OR <input type="radio"/> NOT</div> <div>削除 2013/1-2013/3 PEAK @item</div> <div><input type="radio"/> AND <input checked="" type="radio"/> OR <input type="radio"/> NOT</div> <div>削除 2012/1-2012/3 PEAK @item</div> <div><input checked="" type="radio"/> AND <input type="radio"/> OR <input type="radio"/> NOT</div> <div>削除 2013/1-2013/3 MAX @item</div> <div>実行</div> | <div>削除 2012/1-2012/3 MAX @item</div> <div><input checked="" type="radio"/> AND <input type="radio"/> OR <input type="radio"/> NOT</div> <div>削除 2013/1-2013/3 PEAK @item</div> <div>実行</div> |
|--|---|

図 20. 検索フロー5 の⑤実行後のページ

図 19. 検索フロー5 の②実行後のページ

このような検索フローで検索を行った結果が図 21 になります。  
 クエリの修正を行うことで、同一クエリ間で異なる論理演算子をつなぐことや、不必要になったクエリを削除することが簡単になります。

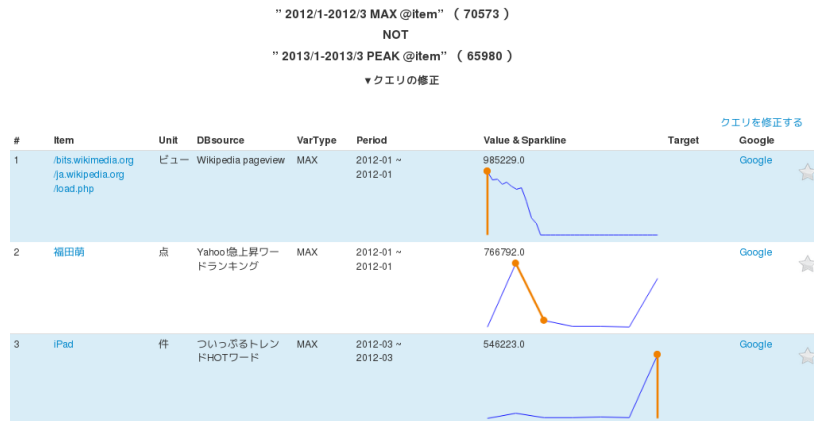


図 21. 検索フロー5 の実行結果(上位 3 件)

## 付録 B 効率的な検索行動に対する知識

3.3 節で作成し、4.3 節の評価実験で用いた効率的な検索行動に対する知識を以下に示す。

### コンテキスト検索エンジンを使う上での注意点

1. @period 検索はアイテム名、@item 検索は期間で行われる。  
検索結果が「@period」の場合は、アイテム名が必須の検索であり、そのアイテムで特徴的変動が発生した期間を結果として表示します。  
検索結果が「@item」の場合は、期間又はアイテム名が必須の検索であり、期間であればその期間の中で特徴的変動が発生したアイテムを結果として表示します。  
アイテム名で検索した場合は、そのアイテムで特徴的変動が発生した期間に同様又は逆の特徴的変動が発生したアイテムを結果として表示します。
2. アイテム名 + @item 検索を行う際には、期間も指定する。  
コンテキスト検索エンジンでの「アイテム名 + @item」は、そのアイテムで特徴的変動が発生した時期に対して検索を行います。  
例 1. 「バレンタインデー + @item + PEAK」で検索を行う  
➡ バレンタインデーは毎年 2 月に PEAK を迎えるので、1 回でも 2 月に PEAK があるアイテムがすべてヒットしてしまう(検索結果が多くなりすぎてしまう)  
なので、「アイテム名 + @item」を行う際には、そのアイテムの動向に対して興味のある期間を選択する必要がある(できれば、1 年程度に期間を絞る)
3. 特徴的変動「MAX、MIN」は AND 検索で複数個つなげる事ができない。  
特徴的変動「MAX、MIN」は各アイテムで動向情報が最も高かった又は低かった 1 点のみを検索を行います。他の PEAK 等の特徴的変動はアイテムの動向に応じて複数個あることが多いが、MAX、MIN は各アイテムに一つしかないので AND 検索で繋げることが出来ない。  
例 2. 「菅直人 + @item + MAX」 AND 「鳩山由紀夫 + @item + MAX」  
➡ 菅直人は 2010 年 6 月に MAX、鳩山由紀夫は 2009 年 9 月に MAX が発生している。  
二つの時期で MAX を持っているアイテムは存在しないので、検索結果が 0 件になる。
4. アイテム名の部分一致は@period 検索でのみ利用する。  
アイテム名の部分一致は、入力されたアイテム名が含まれているアイテム全てを検索結果とします。  
なので、@period 検索では、入力されたアイテム名が含まれているアイテムで特徴的変動が発生した時期を結果として表示します。  
例 3. 「菅直人 (部分一致) + @period + MAX」で検索を行う。  
➡ 「菅直人」の他に「菅直人内閣」や「小菅直人」等 6 つのアイテムで MAX が発生した期間が表示される。  
しかし、@item 検索ではそれら全てのアイテムで特徴的変動が発生した期間で検索を行う。  
例 4. 「菅直人 (部分一致) + @item + MAX」で検索を行う。  
➡ 例 3 の検索で 6 つのアイテムで MAX が発生しており、6 つの期間で MAX が発生している全てのアイテムが検索結果として表示される。(「小菅直人」の MAX 等も含まれてしまう)

## 動向情報に対する知識

1. 動向情報が同じだからといって必ずしも関係があるとは限らない。

動向情報とは、その時の事件や出来事・季節など様々の要因によって決定されます。

なので、ある期間での特徴的変動が一致したからといって必ずしも関係があるとは限らない。

例 1. 「2012 年 1 月 + @item + MAX」の検索結果上位に「平清盛」と「厄年」があった。

➡ 「平清盛」は 2012 年 1 月に大河ドラマ「平清盛」が始まったことが要因と考えられる。

➡ 「厄年」は新年になり、厄年に対しての関心が高くなることが要因と考えられる。

(厄年は毎年 1 月に PEAK を迎える周期性のあるアイテム)

このように、必ずしもアイテム間に関係があるわけではないので、各アイテムがどのような理由でその期間に特徴的変動が発生したのか理解する必要がある。しかし、複数の特徴的変動が一致している場合は、何か同一の要因に影響を受けているか、アイテム間で関連がある可能性が高いです。

2. 期間や特徴的変動だけでなく、動向情報の値も確認する。

現在、コンテキスト検索エンジンで検索できるアイテムの大半が Wikipedia のページビューです。

検索結果の画面から見ることでできる動向情報の概要では、特徴的変動が発生しているがアイテムページに移動して、確認すると全体の値がとても小さいアイテムが存在する。

例 2. 2012 年 1 月に MAX が発生しているアイテムの中に「多良見町」(長崎県の町)がある。

➡ このアイテムの動向情報は MAX が 430、MIN が 215 であった。つまりこのページは多くても月に 430 回しか見られていない。

このような値の小さいアイテムは何かの影響を受けたり、関連があるとは考えにくく特徴的変動も誤差の可能性もある(一般名詞や土地名・あまり有名ではない人名などに多くみられる)。そして、動向情報の値が大きいアイテムの方が、より多くの人が関心を持っているので原因や関係性が特定しやすいです。(値が小さいからといって必ずしも関係性がないというわけでもないので注意が必要である)。

## コンテキスト検索エンジンでの検索のコツ

### 1. 検索したいアイテム・出来事に対する動向情報を理解する。

あるアイテムに関連するアイテムは似たような動向情報を持つことが多い。なので、最初にアイテム名で@period 検索を行い、そのアイテムの動向情報を確認、どの時期にどの特徴的変動が発生したかを理解し、どのような検索クエリを組み合わせればそのアイテムの動向情報を表現できるか考える必要がある。

出来事の場合も同様に、その出来事が起きた時期をしっかりと理解し、検索クエリとして変換する必要がある。

### 2. 検索結果の件数を意識する。

検索を行うと、実行した検索クエリの横に検索結果の件数が表示される。この検索結果の件数があまりにも多いと、適切なクエリで結果を絞り込めていないことが分かる。

例 1. 「2011 年 1 月～2011 年 12 月 +@item + PEAK」で検索を行う。

➡ 検索結果が 1620228 件ととても多く、うまく絞り込めていないと言え難い。

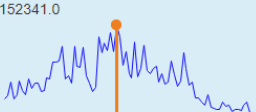
このように、広い期間での検索やアイテム名だけでの検索は検索件数が大きくなりがちなので、複数クエリを組み合わせたり、期間を狭くすることで検索結果を減らす努力が必要である。

### 3. 検索結果を詳しく確認するのは検索結果がある程度絞り込めてからにする。

検索結果が 1000000 件の場合、すべてのアイテムと関係性を調べるのは不可能である。検索結果はランキングされているが、これだけ多い検索件数だと関連のあるアイテムが上位に来ていない可能性もある。なので、検索結果が絞れるまではどのようなアイテムが上位に来ているか確認する程度にとどめ、10000 件以下（出来れば 1000 件程度）まで絞り込んでから、検索結果を詳しく確認した方がよい。

### 4. 知らないアイテム、よくわからないアイテムは一度確認する。

検索結果を見ていく中で知らないアイテムが出てくることがある。知っているアイテムはアイテム間の関係性を理解することができるが、知らないアイテムでは判断することができない。なので、検索結果に知らないアイテムがあった場合は、一度 Google 検索でどのようなものなのか調べてからアイテム間の関係性を調べるとよい。(検索結果の Google ボタンでそのアイテムの Google 検索ができる)

| # | Item       | Unit | DBsource              | VarType | Period               | Value & Sparkline  | Target | Google                 |
|---|------------|------|-----------------------|---------|----------------------|--|--------|------------------------|
| 1 | モンキー・D・ルフィ | ビュ   | Wikipedia<br>pageview | MAX     | 2011-02 ~<br>2011-02 | 152341.0<br> |        | <a href="#">Google</a> |

↑  
これ

## 付録 C 実験に用いた記事

4.2 節および 4.3 節の実験で用いた記事を以下に示す。

◎既成政党への不信増幅＝迷走続きの 3 年－民主政権【12 衆院選】

※記事などの内容は 2012 年 11 月 24 日掲載時のものです

2009 年 8 月末の衆院選で自民党を破り、政権交代を果たした民主党。3 年余りの政権運営は迷走の連続で、首相はこの間に 2 回も代わった。民主党政権への期待は高かっただけに、既成政党に対する有権者の不信が増幅。結果として「第三極」を目指す政党の乱立を招いた。（肩書はいずれも当時）

◇日米同盟損ねる

「脱官僚依存の政治を実践し、無駄遣いを一掃する。政治主導、国民主権、真の意味での地域主権の世の中をつくり上げていく」。09 年 9 月 16 日の鳩山由紀夫首相の就任記者会見は、政権交代の高揚感で満ちていた。

各種世論調査で鳩山内閣発足時の支持率は軒並み 6 割を超え、党運営の実権をあずけた小沢一郎幹事長との「小鳩」体制は、強い追い風を受けてスタートした。

しかし、年末の 10 年度予算編成は財源確保が難航、早くも試練となった。財源対策として小沢氏の主導でマニフェスト（政権公約）の柱の一つだったガソリン税の暫定税率撤廃が見送られたが、子ども手当や高校無償化などを盛り込んだ結果、一般会計総額は初めて 90 兆円を突破。予算の組み替えや無駄削減で 13 年度までに 16.8 兆円の財源を捻出するとしていたマニフェストの弱点が浮き彫りになった。

外交の未熟ぶりはより深刻だった。米軍普天間飛行場の移設問題で、鳩山氏は政権交代前の「最低でも県外」との自らの発言に執着。09 年 11 月、オバマ大統領に「トラスト・ミー（私を信じて）」と早期決着を約束しながら、調整は難航。結局、新たな移設先は見つけられず、自民政権時代に固まっていた名護市辺野古に回帰した。鳩山政権は米国の信頼を失い、「県外」に一時沸いた沖縄県民にも見放される形となった。

小沢氏の元秘書が政治資金をめぐる事件で逮捕され、小鳩の求心力は一気に低下。迫る参院選への影響を懸念する声が民主党内で強まり、鳩山氏は 10 年 6 月、小沢氏を道連れに辞任した。

◇震災対応が混乱

後を継いだ菅直人首相は、7 月の参院選前に唐突に消費増税を表明したが、低所得者対策などをめぐる発言のぶれが批判を浴び、参院選は大敗。「ねじれ国会」が再現した。

9 月、尖閣諸島沖での漁船衝突事件で中国人船長を逮捕・送検しながら、中国側が対抗措置を強めると、船長を釈放。ちぐはぐな対応は野党の追及の的となった。

内閣支持率がじりじりと下がる中、菅氏は政権浮揚を狙って「脱小沢」に動いた。11 年 1 月、小沢氏が政治資金規正法違反罪で強制起訴されると、自発的離党を要求。小沢氏との距離をめぐって党内対立が深刻化した。

3月に発生した東日本大震災は、菅政権を根底から揺さぶった。菅氏をはじめ首相官邸は東京電力福島第1原発事故の対応に追われ、政府内の指揮命令は混乱。被災者支援や復旧・復興の遅れが指摘された。

「菅降ろし」の機会をうかがっていた小沢氏は、野党提出の内閣不信任決議案に賛成する構えを見せ、党内で同調者を募った。このため、菅氏は6月、不信任案の採決に先立ち、震災対応に「一定のめど」を付けた後の退陣を表明せざるを得なかった。8月の特例公債法などの成立を受け、菅氏は退陣した。

#### ◇民自公で消費増税

民主政権で3人目の首相となった野田佳彦氏は就任当初、「親小沢VS反小沢」で繰り返される党内対立に「ノーサイド」を訴えた。小沢氏に近い輿石東氏の幹事長起用は党内融和を重視する決意の表れだった。

だが、野田氏が消費増税路線を鮮明にすると、小沢氏のグループが「マニフェスト違反」との主張を強め、党内抗争が再燃した。

野田氏は、消費増税に政治生命を懸けると明言して退路を断ち、自民、公明両党と連携する路線に傾斜。「近いうち」の衆院解散と引き換えに自公の協力を取り付け、消費増税法を成立させた。同法の採決で小沢氏をはじめ多くの議員が造反。離党者が相次ぎ、民主党は溶解状態に陥った。

外交では野田氏も苦しんだ。9月の尖閣諸島国有化をきっかけに中国各地で反日デモが発生。日中関係は、国交正常化以降で最悪といわれるほど深刻な事態を招いた。日本製品に対する不買運動で、日本経済にも影響が出ている。李明博大統領による竹島上陸も外交力の限界を印象付けた。「第三極」を目指す日本維新の会の石原慎太郎代表は、民主党外交を「弱腰」と批判している。

「政権交代前に時計の針を戻して古い政治に戻るのかどうか問われる選挙だ」。衆院解散後、野田氏はこう叫んだ。しかし、有権者の間には今、民主党政権の継続でも、自民政権への回帰でもない「新しい政権」を求める声が広がっている。

(URL: [https://www.jiji.com/jc/graphics?p=ve\\_pol\\_election-syugin20121124j-04-w680](https://www.jiji.com/jc/graphics?p=ve_pol_election-syugin20121124j-04-w680))